



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta elektrotechnická

Změny v akustických charakteristikách řeči související s věkem

Age-related changes in acoustic characteristics of speech

Diplomová práce

Studijní program: **Biomedicínské inženýrství a informatika**

Studijní obor: **Biomedicínské inženýrství**

Vedoucí práce: **doc. Ing. Roman Čmejla, CSc.**

Praha 2015

Bc. Jarmila Behenská

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: Bc. Jarmila B e h e n s k á
Studijní program: Biomedicínské inženýrství a informatika (magisterský)
Obor: Biomedicínské inženýrství
Název tématu: Změny v akustických charakteristikách řeči související s věkem

Pokyny pro vypracování:

Cílem diplomové práce je nalezení a ověření akustických charakteristik charakterizujících věk. Na základě rešerše literatury vyberte a vyhodnoťte vhodné akustické parametry závislé na věku. Zaměřte se také na vyhodnocení věkové závislosti akustického parametru počátek znělosti (Voice Onset Time). Proveďte rozbor možností vytvoření databáze pro studium věkově závislých akusticko-fonetických jevů. Na části databáze proveďte akusticko-fonetické analýzy se zaměřením na věkovou závislost.

Seznam odborné literatury:

- [1] Košturiaková, L.: Věková závislost lidské řeči. Diplomová práce, ČVUT FEL, Praha 2013.
- [2] Janda, J.: Posuzování logopedického věku dítěte. Disertační práce, ČVUT FEL, Praha 2013.
- [3] Reubold, U., Harrington, J., Kleber, F.: Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers. *Speech Communication* 52, 638–651, 2010.
- [4] Harrington, J., Palethorpe, S., Watson, C.I.: Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers. *Interspeech*, 2, pp.1081–1084, 2007.
- [5] Schötz, S.: Acoustic Analysis of Adult Speaker Age. *Lecture Notes in Computer Science*, 4343 LNAI, pp. 88–107, 2007.
- [6] Psutka, J., Matoušek, J., Müller, L., Radová, V.: Mluvíme s počítačem česky, Academia, 2006.

Vedoucí diplomové práce: doc. Ing. Roman Čmejla, CSc.

Platnost zadání: do konce zimního semestru 2015/2016

L.S.

prof. Ing. Pavel Sovka, CSc.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 15. 9. 2014

Abstrakt

V dnešní době je důležitý výzkum věkové závislosti změn hlasových charakteristik pro včasné rozpoznání změn, které nesouvisí s fyziologickým stárnutím zdravého hlasu. Charakteristiky změn řečových signálů mají své uplatnění v syntézách řeči, metodách rozpoznávání osob, ale i komerčních aplikacích pracujících s ovládáním hlasem.

Experimentální ověřování věkové závislosti akustických parametrů hlasu proběhlo v této práci na databázi vytvořené z filmových záznamů herečky Ivany Chýlkové v jejím z období 1985 – 2014. Data byla převedena z formátu avi na wav a ze všech záznamů byly nastříhány různě dlouhé promluvy. Zpracování zvukových signálů a následná analýza vybraných akustických parametrů byla provedena v Praatu, sofistikovaném softwaru, který se standardně používá pro zvukové analýzy. Akustické parametry byly vybrány na základě rešerše dané problematiky.

V této práci byla provedena podélná studie ženského hlasu v rozmezí třiceti let. Byly analyzovány a hodnoceny parametry základní frekvence F0, jitter, shimmer, první dvě formantové frekvence F1 a F2, velikost vokálního prostoru VSA, počátek znělosti VOT a rychlost řeči. Výzkum byl proveden na dlouhých záznamech i na segmentech, slabikách a hláskách.

Abstract

In nowadays science field, the research of age-related changes in acoustic characteristics of speech is dependent for early detection changes, that are not related to physiological aging of healthy voice. The characteristics of changes in speech have their importance in speech synthesis, speech recognition methods and in commercial applications, which work with voice control.

Vocal aging dependence of acoustic parameters of voice was experimental verification on the database. The database was created from films records of the actress Ivana Chýlková in the period from 1985 to 2014. This data was converted from AVI to WAV format. Audio signal processing and the vocal analysis of selected acoustic parameters was processed in sophisticated software Praat, which is used for measuring standard of acoustic signals. The acoustic parameters were selected based on the research of the issue.

This diploma thesis deals with the longitudinal study of the women's voice within thirty years. The most appropriate appeared the evaluation of the fundamental frequency F0, jitter, shimmer, the first and the second formant, vocal space area, voice onset time and speech rate. The research was evaluated on long records, on segments of speech and syllables.

Klíčová slova

Akustická analýza hlasu, podélná studie, Praat, hlas, řeč, věkově závislé změny, základní frekvence, formantové frekvence, počátek znělosti, jitter, shimmer, vokální trojúhelník, artikulační index.

Keywords

Acoustic analysis of voice, longitudinal study, Praat, voice, speech, age – related changes, fundamental frequency, formants, voice onset time, jitter, shimmer, vowel space area, vowel articulation index.

Poděkování

Ráda bych poděkovala svému vedoucímu diplomové práce, panu doc. Ing. R. Čmejlovi, CSc., za veškerou pomoc, za čas, který mi věnoval, trpělivost, cenné rady a připomínky při vedení mé diplomové práce. Dále bych ráda poděkovala své rodině a přátelům za morální a finanční podporu při studiu.

Čestné prohlášení

Prohlašuji, že jsem diplomovou práci s názvem Změny v akustických charakteristikách řeči související s věkem vypracovala samostatně a použila k tomu úplný výčet citací použitých pramenů, které uvádím v seznamu přiloženém k práci.

Nemám závažný důvod proti použití tohoto školního díla ve smyslu §60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne 11. května 2015

.....

Behenská Jarmila

1. Obsah

Abstrakt	ii
Klíčová slova	iii
Poděkování	iv
Čestné prohlášení.....	v
Seznam obrázků.....	viii
Seznam tabulek.....	ix
Seznam zkratk.....	x
1. Obsah	vi
2. Úvod	1
3. Analýza zkoumané problematiky	2
3.1 Podélné studie	2
3.1.1 Přehled literatury	2
3.1.2 Přehled výsledků podélných studií.....	6
3.2 Příčné studie.....	9
3.2.1 Přehled literatury	9
3.2.2 Přehled výsledků příčných studií.....	14
4. Lidský hlas	19
4.1 Artikulace.....	20
4.1.1 Samohlásky.....	20
4.1.2 Souhlásky	22
4.1.2.1 Explozívy	22
5. Praat	24
5.1 Možnosti Praatu	24
6. Věkově závislé parametry	26
6.1 Fonační vlastnosti řeči	26
6.1.1 Základní frekvence	26
6.1.2 Frekvenční variabilita jitter	27
6.1.3 Amplitudová variabilita shimmer.....	28

6.2 Artikulační vlastnosti řeči	28
6.2.1 Formantové frekvence	29
6.2.2 VSA	29
6.2.3 VAI	29
6.2.4 F2i/F2u	30
6.2.5 Počátek znělosti VOT	30
6.2 Prozodické vlastnosti řeči	31
6.3.1 Rychlost řeči	32
6.3.2 Podíl pauz v promluvě	32
7. Analýza věkové závislosti	33
7.1 Databáze	33
7.2 Základní frekvence	34
7.3 Jitter a shimmer	35
7.3.1 Hodnocení variabilit F0	36
7.4 Formantové frekvence	37
7.4.1 Hodnocení formantů F1, F2	38
7.5 Vokalický trojúhelník a VSA	39
7.6 VAI a poměr F2i/F2u	40
7.6.1 Hodnocení VAI a F2i/F2u	41
7.7 Počátek znělosti VOT	41
7.7.1 Hodnocení VOT	44
7.8 Rychlost řeči	45
7.8.1 Hodnocení rychlosti řeči	45
7.9 Zhodnocení výsledků	46
8. Závěr	47
9. Naplnění cílů, perspektivy dalšího řešení	48
10. Seznam použité literatury	49
A Obrazová příloha	51
B Tabulková příloha	58
C Seznam filmů	60
D Obsah příloženého CD	61

Seznam obrázků

3.1	Trendy F0, F1, F2, F3 s rostoucím věkem pro ženský hlas	4
3.2	Změny formantů pro jednotlivé vokály	4
3.3	Změny F0 po 25 let u 8 ženských subjektů	5
3.4	Průběh F0 a F1, vlevo Elisabeth II, vpravo mužský hlas	5
4.1	Artikulační orgány	19
4.2	Schéma vzniku hlasu podle teorie zdroje a filtru	20
4.3	Hellwagův trojúhelník	20
5.1	Editor okno Praat	24
5.2	Editor okno s textovým popisem	25
6.1	Frekvenční variabilita	27
6.2	Amplitudová variabilita	28
6.3	Sekvence VEV a segmenty	30
7.1	Průběh F0 s věkem	34
7.2	Průběh jitteru s věkem	35
7.3	Průběh shimmeru s věkem	36
7.4	Průměrné hodnoty F1	37
7.5	Průměrné hodnoty F2	38
7.6	Změny pozic samohlásek s věkem	39
7.7	Průběh VSA s věkem	39
7.8	Vokalické trojúhelníky ve 22 a 51 letech	40
7.9	Průběh parametru VAI s věkem	40
7.10	Průběh poměru F2i/F2u s věkem	41
7.11	Průběh doby exploze (VOT)	43
7.12	Průběh doby okluze	43
A.1-A.19	VOT pro jednotlivé roky /u-dʰ-e/	51

Seznam tabulek

3.1	Trendy F0, F1, F2, F3 s věkem pro ženský hlas	4
3.2	Trendy F0, F1, F2, F3 s věkem pro mužský hlas	4
3.3	Výsledné trendy s věkem	6
3.1.1	Shrnutí výsledků podélných studií	6
3.4	Trendy F0, jitteru, shimmeru s věkem	9
3.5	Počáteční hypotézy změn akustických parametrů	10
3.6	Výsledky naměřených změn frekvencí na vokálu	11
3.7	Výsledky trendů akustických parametrů s věkem	12
3.8	Doby trvání exploziv /d'/	13
3.9	Doby trvání exploziv /d'/ od jiných autorů	13
3.10	Trendy akustických parametrů Jandovy studie	14
3.2.1	Shrnutí výsledků příčných studií	14
4.1	Hodnoty F1 a F2 pro české samohlásky	21
4.2	Rozdíly při tvorbě českých znělých a neznělých exploziv	23
6.1	Rozsah F0 řeči	26
7.1	Databáze extrahovaných záznamů z filmů herečky	33
7.2	Počet analyzovaných prodloužených fonací /a/	34
7.3	Počet hodnot pro výpočet jitteru	35
7.4	Počet analyzovaných prodloužených fonací /a/, /e/, /i/, /o/, /u/	37
7.5	Naměřené doby trvání /u-d'-e/	42
7.6	Naměřené doby trvání /e-d'-e/	44
7.7	Naměřené doby trvání slabik	45
7.8	Výsledky věkové závislosti akustických parametrů	46
B.1	Průměrné hodnoty F0 pro /a/	58
B.2	Průměrné hodnoty jitteru, shimmeru	58
B.3	Průměrné hodnoty F1, F2 pro /a/, /e/, /i/, /o/, /u/	59
B.4	Průměrné hodnoty dob exploze, okluze	59
C.1	Seznam filmů databáze	60

Seznam zkratek

F0	Základní frekvence
SFF	Základní frekvence
SFFsd	Směrodatná odchylka základní frekvence
Jitter	Frekvenční variabilita
Shimmer	Amplitudová variabilita
F1 – F5	Formantové frekvence
VOT	Voice onset time, počátek znělosti
VSA	Vocal space are, velikost vokální plochy
VAI	Vowel articulation index, artikulační index
F2i/F2u	Poměr formantových frekvencí pro samohlásky /i/, /u/
CVC	Konsonanta – vokál – konsonanta (souhláska – samohláska - souhláska)
VCV	Vokál – konsonanta – vokál (samohláska – souhláska - samohláska)
VEV	Vokál – explozíva – vokál

2. Úvod

Téma závěrečné práce s názvem Změny v akustických charakteristikách řeči související s věkem mě zaujalo z důvodu možnosti včasné diagnózy neurodegenerativních chorob, aplikace získaných trendů zdravého jedince pro vývoj akustických analýz a metod rozpoznávání řeči pro různé věkové skupiny, pro rozvoj komerčních aplikací pracujících s hlasem a pro celkový náhled na fyziologické, anatomické i lingvistické důsledky vlivem stárnutí hlasu.

Základní vlastnosti lidského hlasu jsou dány fyziologií hlasového ústrojí. Je známo, že vlivem stárnutí dochází k fyziologickým změnám organismu a tak i samotných akustických parametrů hlasu. Změny výšky hlasu vznikají změnami postavení hlasivkových chrupavek, proměnou napětí hlasivek, změnou tlaku výdechového vzduchu i práce svalů hrtanu. Hlas vycházející z hlasivek získává barvu průchodem jednotlivými rezonátory, typickými pro každého jednotlivce. S vyšším věkem se mění i vlastnosti rezonančních dutin a tím dochází ke změnám rezonančních frekvencí neboli formantů.

Věkovou závislostí některých hlasových parametrů se zabývá řada příčných i podélných studií, na jejichž základě jsem vybrala pro moji práci vhodné akustické parametry. Studie zkoumaly jak mužské, tak ženské hlasy. Příčné studie zkoumají hodnoty u více jedinců v každé věkové kategorii. Sběr dat od různých věkových skupin způsobuje zkreslení průběhů parametrů vlivem individuální variability. Podélné se zabývají studiem jediného jedince během života. Záznam hlasu jedince v průběhu života nelze prakticky získat pro rychlou analýzu. Jako vhodné řešení se nabízí využít rozhlasové nebo filmové záznamy herců pro analýzu průběhů hlasových charakteristik.

Metodologie akustických analýz je postavena zkoumání třech základních vlastností řečového signálu, na fonačních, artikulačních a prozodických vlastnostech řeči. Fonační analýza zkoumá průměrné hodnoty a odchylky základní frekvence, jitteru, shimmeru. Artikulační se zabývá vokálním prostorem a artikulačním indexem, což jsou příznaky, podle kterých se dají rozpoznat nemocní od zdravých. Prozodické řeší rychlost artikulace a pauzy v promluvách.

V práci jsou analyzovány filmové a rozhlasové nahrávky herečky Ivany Chýlkové v období 1985 – 2014. V rámci návrhu samotného hodnocení z pohledu fonetického, artikulačního i prozodického jsem vybrala vhodné akusticko fonetické parametry hlasu. Jde o základní frekvenci, první dva formanty, vokální trojúhelník, počátek znělosti, jitter, shimmer, artikulační rychlost.

3. Analýza zkoumané problematiky

Se stárnutím lidské řeči dochází ke změnám různých hlasových parametrů. Řada studií se zabývala hodnocením těchto změn. Jejich cílem bylo vysvětlit důsledky normálního stárnutí na řečové charakteristiky. V jednotlivých studiích byly použité různé metodiky výzkumu, zabývaly se různými akusticko fonetickými jevy, zkoumaly vliv věku i pohlaví na hlasový signál. Výsledky dosavadních studií se částečně shodují, ale částečně jsou v rozporu. Neexistuje tedy zatím úplné objasnění změn hlasových parametrů způsobených změnou věku. V této práci je uvedena rešerše studií zabývajících se problematikou věkové závislosti. Jsou zde popsány výsledky trendů pro jednotlivé zkoumané parametry. Na základě rešerše byla navržena metodika analýzy věkově závislých řečových charakteristik a ověřena na databázi.

3.1 Podélné studie

Podélné studie sledují průběhy změn hlasových parametrů během života jedince. Databáze zkoumaných objektů se zpravidla pořizuje po určitém časovém intervalu. Kvalita použitých záznamů se liší v použité technice nahrávání. Podélné výzkumy jsou založeny na archivních záznamech. Výsledky popisují trendy analyzovaných řečových charakteristik pro daný subjekt. U pohlaví se vyskytují rozdílné výsledky. Cílem je provést co nejvíce podélných studií na různých subjektech a zobecnit výsledné trendy pro celou populaci žen a mužů. Některé studie mohou mít zcela intraindividuální výsledky, které se nedají zobecnit.

Tato práce se zabývá analýzou ženského hlasu. V rešerši jsou uvedené zkoumané parametry pro ženské i mužské hlasy a zobecněné výsledky v tabulce 3.1.1.

3.1.1 Přehled literatury

Decoster W. a Debruyne F. analyzovali ve studii [1] trend průměrné hodnoty základního hlasivkového tónu (SFF), standardní odchylku (SFFsd) a počátek znělosti (VOT). Parametry byly zkoumány na archivních záznamech, na páskách. Z toho důvodu byly z analýzy vyřazeny parametry jitter a shimmer, které by mohly být deformované kvalitou pásky.

Byly pořízené dvě nahrávky v rozmezí třiceti let. Tuto dobu považovali za dostatečně dlouhý časový úsek pro očekávání objevení hlasových změn.

Měřenými subjekty bylo dvacet mužů, novinářů. První nahrávky byly pořízeny ve věku 29 – 32 let. Po třiceti letech došlo ke změně zdravotního stavu subjektů, pouze 7 mužů zůstalo v aktivní profesi. Po této době nebylo dosaženo ekvivalentního hlasového zatížení.

SFF a SFFsd byly analyzovány z dlouhých vět bez dlouhých pauz a přerušení. Průměrná doba trvání prvního archivního záznamu bylo 11,86 s (rozsah 5,5 – 16 s), doba druhého se pohybovala kolem průměrné hodnoty 12,06 s (rozsah 6,8 – 15,56 s).

VOT byl analyzován ze sekvence souhláska – samohláska a stanoven jako doba mezi uvolněním souhlásky a prvním glottálním pulsem následující samohlásky. Hraniční body byly určeny ručně ve spektrogramu. Ze záznamů byly vybírány slabiky /pa/, /ka/ z různých slov.

Výsledkem práce byl klesající trend základní frekvence u mužů s rostoucím věkem. SFF klesla ze 138 Hz na 125 Hz během třiceti let. Standardní odchylka byla hodnocena také jako klesající s věkem, a sice změna z 26,2 Hz na 21,9 Hz. VOT naopak u mužů s rostoucím věkem vzrostl, pro slabiku /pa/ z 13,7 ms na 19,5 ms a pro slabiku /ka/ z 22,2 ms na 31,5 ms.

Ve studii [2] z roku 2007 byla provedena experimentální podélná studie ženského hlasu. Zkoumali, zda časový úsek pěti let přinese významné změny v hlasu. Z hlasových parametrů byly analyzovány základní frekvence, první tři formantové frekvence samohlásek /a/, /i/, /u/, rychlost řeči (počet slabik za sekundu ve větě), pauzy v promluvě.

Pro studii bylo vybráno 9 žen a časový úsek nahrávek pět let. Nahrávky se skládaly ze čtené i spontánní mluvy. Ženy měly za úkol přečíst tři věty o popisu cesty. Spontánně měly mluvit o obrázku, který dostaly. Řečové vzorky byly dále manuálně segmentovány a zpracovávány pomocí softwaru Praat.

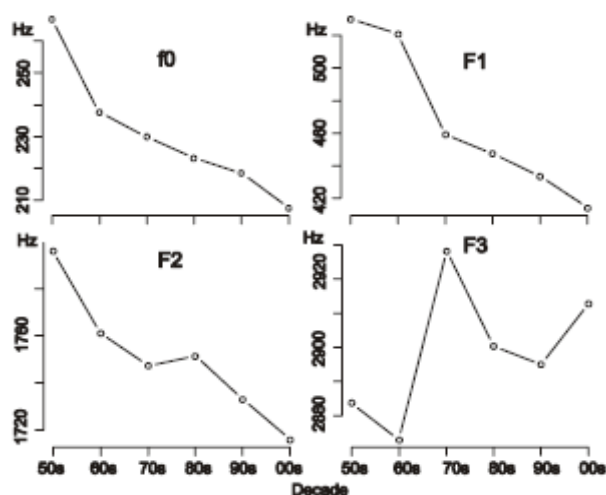
Z výsledků studie jsou patrné pouze fakty, že F0 a F1 jsou dobrými indikátory věku a s rostoucím věkem klesají. Průběh F0 byl sledován na větách spontánní řeči. Po pěti letech byly hodnoty nižší. Ostatní formantové frekvence nejsou označené za významné ukazatele věku. Tempo řeči je hodnoceno jako nejspolehlivější akustický ukazatel věku ze spontánní řeči. S rostoucím věkem je tempo pomalejší, ale nejsou ukázány ani hodnoty, ani průběh.

Ve studii [3] také z roku 2007 byla zkoumána věková závislost vybraných akustických parametrů na dvou ženských a na dvou mužských hlasech. Analyzována byla základní frekvence, první tři formanty, vokálníky trojúhelník, změny v rozdílech F1/F0 a F3/F2.

Databáze pro analýzu ženského hlasu královny Elisabeth II byla vytvořena z vánočních vystoupení v každé dekádě od roku 1950 do roku 2000, od věku 24 do 74 let. Průměrná doba nahrávek činila 5,2 minuty (rozmezí 1 – 8,2 minut). Záznamy druhého ženského hlasu byly pořízeny z rozhovorů s Marií Lockwood v radiových pořadech v časovém úseku třiceti let. První záznam byl délky 5,5 min, kdy jí bylo 35 let, a druhý 12 min dlouhý, kdy jí bylo 64 let. Databáze mužských hlasů byla vytvořena v časovém intervalu čtyřiceti let opět v radiových pořadech o délce zhruba dvaceti minut. První mužský hlas byl analyzován ve věku 40 a 79 let. Druhý ve věku 36 a 73 let.

F0 i všechny formantové frekvence byly analyzovány na znělých segmentech vět, tedy na samohláskách. F1 má v období šesti dekád klesající trend, F2 až na mírnou změnu mezi 44. a 54. rokem života královny má také klesající trend. Třetí formant má skokové změny. Výsledky této studie pro ženský hlas královny je ukázán na obr. 3.1.

U druhého ženského hlasu byl potvrzen klesající trend F0 a F1 s rostoucím věkem. V tab. 3.1 jsou uvedené výsledky změn.



Obr.3.1: Výsledné trendy F0, F1, F2, F3 s rostoucím věkem pro ženský hlas [3].

(35 let)	F0 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]	(64 let)	F0 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]
Mladší	156	481	1508	2265	Starší	123	474	1503	2225

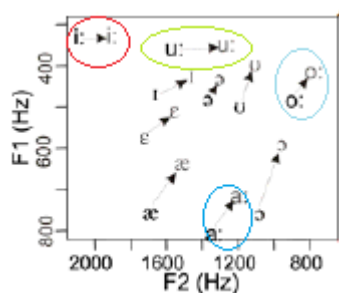
Tab. 3.1: Výsledné trendy F0, F1, F2, F3 s rostoucím věkem pro ženský hlas [3].

Výsledky pro mužské hlasy jsou uvedené v tab. 3.2. F0, F1 i F2 klesají u obou. F3 u jednoho klesá, u druhého roste s věkem.

1.muž	F0 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]	2.muž	F0 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]
Mladší (40 let)	166	535	1578	2510	Mladší (36 let)	131	479	1320	2265
Starší (79 let)	123	420	1429	2594	Starší (73 let)	111	429	1259	2225

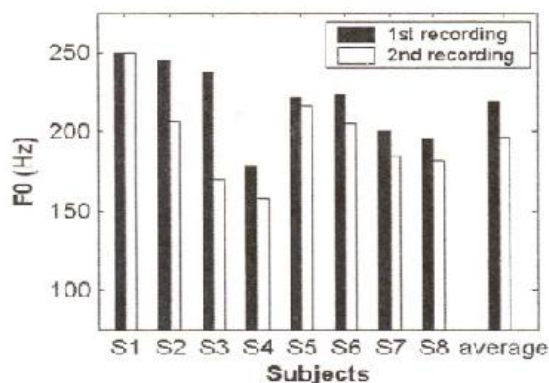
Tab. 3.2: Výsledné trendy F0, F1, F2, F3 s rostoucím věkem pro mužský hlas [3].

Dále byly hodnoceny změny kombinací formantů F1 a F2 samohlásek ve vokalizickém trojúhelníku. Výsledky změn mužského hlasu jsou zobrazeny na obr. 3.2. V obrázku jsou zvýrazněné změny samohlásek, které jsou zajímavé pro analýzu v této práci. Pro prodlouženou fonaci samohlásky /a/ a /o/ klesá F1 i F2. U samohlásek /i/ a /u/ je hodnota F1 stejná a F2 klesá.



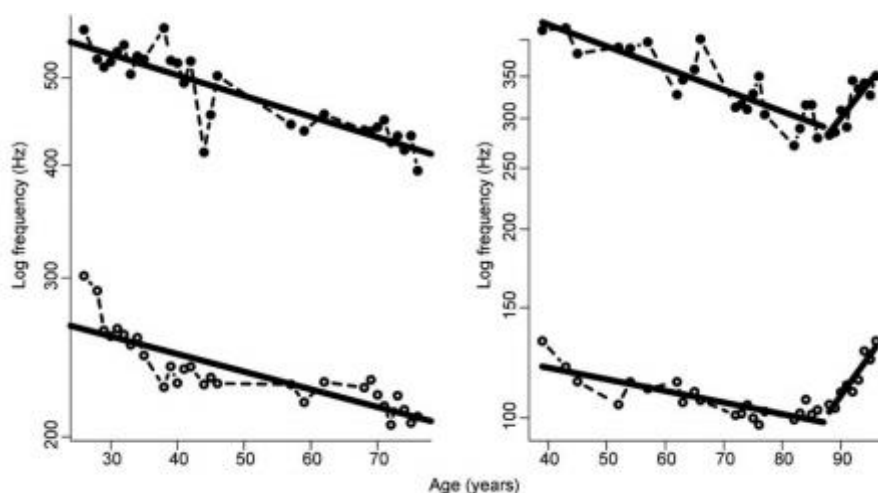
Obr. 3.2: Změny formantů pro jednotlivé vokály [3].

Studie [4] publikovaná v roce 2008 se snažila objasnit změny v základní frekvenci ženského hlasu s rostoucím věkem. Analýza byla provedena na osmi ženských hlasech s časovým intervalem měření 25 let. Zkoumané věkové rozmezí bylo 26 – 38 let při prvním měření, při druhém 51 – 63 let. Všechny ženy byly Američanky s podobným fyzickým stavem. F0 byla analyzována ze záznamu o délce 15 sekund u všech žen. Výsledek studie je ukázán v obr. 3.3. Studie potvrdila významný klesající trend F0 u žen s rostoucím věkem. Z obrázku lze vyčíst maximální naměřenou počáteční hodnotu F0 250 Hz u subjektu 1, minimální je kolem hodnoty 155 Hz u subjektu 4. Ženy jsou přibližně stejného věku, ale přesto je pozorována variabilita v hodnotách F0. Důležitým ukazatelem je klesající trend.



Obr. 3.3: Změny F0 po 25 let u 8 ženských subjektů [4].

Analýza [5] z roku 2010 zkoumala, do jaké míry rostoucí věk ovlivňuje základní frekvenci a první tři formanty. Výzkum byl proveden na dvou získaných nahrávkách z archivů od dvou žen a tří mužů. Oba ženské hlasy byly již analyzovány v předešlé studii [3]. Nahrávky ženských hlasů byly vzorkovány frekvencí 16 kHz, mužských 22 a 24 kHz. V práci se zpracovalo celkem 1064 nahrávek samohlásek, 447 při prvním měření, 617 při druhém. Věková závislost byla prokázána pouze u parametrů F0 a F1. U F2 a F3 nebyla prokázána spojitost s věkem. F0 i F1 klesají s věkem u všech subjektů, u mužů i žen. Dále našli autoři studie vztah mezi F0 a F1. Charakteristika F1 aproximuje průběh F0, viz obr. 3.4.



Obr. 3.4: Průběh F0 a F1, vlevo Elisabeth II, vpravo mužský hlas [5].

Košturiaková ve své diplomové práci [14] provedla podélnou studii akustických parametrů na ženském hlase herečky Ivany Chýlkové. Zkoumané parametry v práci byly F0, F1, F2, jitter, doba trvání exploze (VOT) pro explozivitu /t/, harmonicita, spektrální sklon, rychlost řeči, velikost vokálního prostoru.

F0, F1, F2, jitter byly získány ze slov obsahující prodlouženou fonaci samohlásky /a/. F0 byla naměřena také z celých promluv. Z každého filmu získala data minimálně pro čtyři různé promluvy. Počátek znělosti (VOT) byl zkoumán na extrahovaném slově /dítě/. Rychlost řeči zkoumala na extrahovaném slově /sama/.

Výsledky jejího experimentu jsou uvedené v tab. 3.3.

Parametr	Trend
F ₀ [Hz]	Klesá
F ₁ [Hz]	Klesá
F ₂ [Hz]	pro /a/, /u/ roste, pro /i/ klesá
Jitter	Klesá
NHR	Roste
HNR	Klesá
Autokorelace	Nemění se
Spektrální sklon	Nelze určit trend
Rychlost řeči	Roste
Vokální trojúhelník	Zmenšuje se
Doba exploze /t/	Klesá

Tab.3.3: Výsledné trendy s rostoucím věkem [14].

3.1.2 Přehled výsledků podélných studií

V tab. 3.1.1 jsou uvedené výsledky trendů pro jednotlivé akustické parametry, které byly analyzovány v podélných studiích zmíněných výše v přehledu literatury.

Tab. 3.1.1: Shrnutí výsledků podélných studií.

Základní frekvence F0 [Hz]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	F0 [Hz]	SD [Hz]
[1]	20 mužů (29-32)	30 let (59 – 62)	Radiové záznamy	Klesá 138-125	Klesá 26,2–21,9
[3]	Muž 1 (40)	39 let (79)	Radiové záznamy, znělé segmenty vět	Klesá 166-131	
[3]	Muž 2 (36)	37 let (73)	Radiové záznamy, znělé segmenty vět	Klesá 123-111	

Tab. 3.1.1: Pokračování.

Základní frekvence F0 [Hz]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	F0 [Hz]	SD [Hz]
[5]	Muž 1 (35) Muž 2 (43) Muž 3 (37)	35 let (70) 30 let (73) 34 let (71)	Radiové záznamy	Klesá	
ŽENSKÉ HLASY					
[2]	9 žen (různý věk)	5 let	Čtený projev řeči (3 věty) Spontánní řeč (popis obrázku)	Klesá	
[3]	Elisabeth II. (24)	10 let 6 dekád	Záznamy vánočních proslovů, znělé segmenty vět	Klesá 270-200	
[3]	Marie Lockwood (35)	29 let (64 let)	Radiové záznamy, znělé segmenty vět	Klesá 156-123	
[4]	8 žen (26-38)	25 let (51-63)	15 s záznam hlasu	Klesá 250-155	
[5]	Elisabeth II (34) Marie (35)	34 let (68) 29 let (64)	Záznamy z vánočních vystoupení a radiové záznamy	Klesá	
[14]	Chýlková (20)	28 let (48)	Filmové záznamy Prodloužená fonace /a/	Klesá	
Formantové frekvence F1, F2 [Hz]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	F1 [Hz]	F2 [Hz]
[3]	Muž 1 (40)	39 let (79)	znělé segmenty vět	Klesá 535-420	Klesá 1578-1429
[3]	Muž 2 (36)	37 let (73)	znělé segmenty vět	Klesá 479-429	Klesá 1320-1259
[3]	2 muži (36, 40)	37,39 let (73, 79)	/a/, /o/ /i/, /u/	Klesá Stejná	Klesá Klesá
[5]	Muž 1 (35) Muž 2 (43) Muž 3 (37)	35 let (70) 30 let (73) 34 let (71)	Radiové záznamy	Klesá	-
ŽENSKÉ HLASY					
[2]	9 žen (různý věk)	5 let	Spontánní řeč (popis obrázku), Vokály /a/, /i/, /u/	Klesá	-
[3]	Elisabeth II. (24)	10 let 6 dekád	znělé segmenty vět	Klesá	Klesá

Tab. 3.1.1: Pokračování.

Formantové frekvence F1, F2 [Hz]					
ŽENSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	F1 [Hz]	F2 [Hz]
[3]	Marie L. (35)	29 let (64 let)	znělé segmenty vět	Klesá 481-474	Klesá 1508-1503
[5]	Elisabeth II (34) Marie (35)	34 let (68) 29 let (64)	Záznamy z vánočních vystoupení a radiové záznamy	Klesá	-
[14]	Chýlková (20)	28 let (48)	Filmové záznamy, Prodloužená fonace /a/	Klesá	Roste
Počátek znělosti VOT [ms]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	VOT [ms]	SD [ms]
[1]	20 mužů (29-32)	30 let (59 – 62)	Radiové záznamy Slabiky /pa/, /ka/	Roste 13,7-19,5 22,5-31,5	
ŽENSKÉ HLASY					
[14]	Chýlková (20)	28 let (48 let)	/dítě/ - explozíva /t/	Klesá	
Rychlost řeči [slabik/s]					
ŽENSKÉ HLASY					
[14]	Chýlková (20)	28 let (48 let)	/sama/	Klesá	

Na základě shrnutí výsledků ze zpracovaných podélných studií bylo navrženo provést podélnou analýzu ženského hlasu. Databáze dat by byla vytvořena z filmových záznamů ženy, navzorkovány se stejnou vzorkovací frekvencí, podobné kvality.

Navržené akustické parametry a metodika je:

- Základní frekvence F0 [Hz] z celých vět i ze znělých segmentů záznamů, na prodloužené fonaci samohlásky /a/.
- Formantové frekvence F1, F2 [Hz] na prodloužených fonacích samohlásek /a/, /e/, /i/, /o/, /u/.
- Počátek znělosti VOT [ms] na sekvenci vokál – explozíva - vokál

Předpoklady na základě výsledků z tabulky 3.1.1:

- F0 bude s rostoucím věkem klesat.
- F1 u všech vokálů bude klesat, F2 bude klesat u /a/, /o/

3.2 Příčné studie

Analyzovaným objektem není průběh hlasových parametrů v průběhu života jednoho jedince, ale jde o výzkum hlasových parametrů pro různé věkové skupiny. Dané věkové skupiny vyžadují velké množství naměřených dat, protože je značná intravariabilita hlasu mezi různými jedinci stejného pohlaví.

V přehledu literatury je uvedeno několik příčných studií zkoumající problematiku věkové závislosti hlasových parametrů. Výsledky studií jsou zobrazeny v tab. 3.2.1.

3.2.1 Přehled literatury

Xue, Deliyski (2001) [6] porovnávali 15 akustických parametrů starší věkové skupiny s databází normovaných hodnot mladších mužů i žen. Měřenými subjekty bylo 21 starších mužů (průměrný věk 75,4 let) a 23 starších žen (průměrný věk 74,8 let). Nahrávány byly záznamy fonace samohlásky /a/ delší než 3 sekundy, na kterých byla měřena základní frekvence, lokální jitter, lokální shimmer. Před samotným nahráváním byli měřené subjekty podrobeny artikulačnímu cvičení. Pro srovnání s takto získanými hodnotami vybrali autoři normované hodnoty mladších subjektů, které byly zveřejněny v předchozí práci jednoho z autorů této studie (Deliyski & Gress, 1988). V tab. 3.4 je uveden výsledný trend průměrných hodnot a směrodatných odchylek vybraných akustických parametrů s rostoucím věkem.

	F0 [Hz]		Jitter [%]		Shimmer [%]	
	M	SD	M	SD	M	SD
Muži	Klesá (145/128)	Roste (23/29)	Roste (0,6/2,1)	Roste (0,5/1,6)	Roste (2,5/5,5)	Roste (1,0/3,5)
Ženy	Klesá (244/188)	Roste (27/42)	Roste (0,6/2,0)	Roste (0,3/2,0)	Roste (2,0/5,3)	Roste (0,8/4,5)

Tab. 3.4: Výsledné trendy F0, jitteru, shimmeru s rostoucím věkem, M = průměr, SD směrodatná odchylka [6].

Ve studii (2007) [7] byla vytvořena databáze mužských i ženských hlasových záznamů pro analýzu osmi věkových skupin rozdělených po dekáдах. Jednalo se o věk 20 – 90 let. Analýza parametrů byla zkoumána na šesti izolovaných slovech, která byla vybrána základě předchozí studie. Na celém souboru dat byly analyzovány charakteristiky:

- Rychlost řeči [počet slabik za sekundu],
- délka segmentů [ms],
- úroveň tlaku vzduchu SPL [dB],
- základní frekvence F0 [Hz],
- jitter,
- shimmer,

- formantové frekvence (F1 – F5 [Hz]).

Nejdříve si autoři na základě zpracované rešerše stanovili hypotézy změn akustických parametrů s rostoucím věkem, viz tab. 3.5.

	Délka segmentu [ms]	Rychlost řeči [sl/s]	SPL [dB]	F0 [Hz]	Jitter, shimmer	formanty
Ženy	Roste	Klesá	Roste	Stabilní do 50, pak roste	Roste	Klesají
Muži	roste	Klesá	Roste	Klesá do středního věku, pak roste	Roste	Klesají

Tab. 3.5: Počáteční hypotézy změn akustických parametrů [7].

Výsledné průběhy průměrných normalizovaných hodnot parametrů byly v této studii a jsou uvedeny níže:

- F0 u žen klesá do věku 50ti let, do 60ti je konstantní , pak mírně narůstá do věku 80 let a poté opět klesá.
- F0 u mužů klesá do čtyřicátého roku, pak je konstantní do 50ti, do 80 roste a poté prudce roste.
- Délka segmentů šesti zkoumaných slov se s věkem prodlužuje u obou pohlaví.
- Úroveň akustického tlaku má rostoucí trend pro obě pohlaví.
- Lokální shimmer je u žen do věku 50 konstantní. Mezi 50. a 70. rokem klesá, poté roste.
- Lokální shimmer u mužů do čtyřiceti let roste, pak klesá do 70, zde je stabilní a po 80. roce života začíná klesat.
- F1 u obou pohlaví klesá a F2 roste s rostoucím věkem.

Z výsledků vyplývá, že se nepotvrdila předpokládaná hypotéza u parametrů jitter a shimmer, které v této studii dosahovaly s rostoucím věkem spíše nižších hodnot. Výsledky rychlosti řeči zde nebyly zmíněny, pouze, že se prodlužují segmenty. V práci [8], ze které tento článek vyplývá pojednává o nárůstu rychlosti řeči v souhláskách a explozivech. To znamená, že pokud se zkracuje doba exploze (VOT) a prodlužuje doba segmentů (samohlásek), pak rychlost řeči zůstává konstantní.

Torre, Barlow publikovali v roce 2009 příčnou studii [9] hodnotící vliv věku na základní a první tři formantové frekvence (F0, F1, F2, F3) a počátek znělosti (VOT). Analýza byla zkoumána na 27 mladých (15 žen, 12 mužů průměrného věku 25,5 let) a 59 starších subjektech (32 žen a 27 mužů průměrného věku 75,2 let). Úkolem měřených subjektů bylo přečíst list 22 segmentů souhláska – samohláska – souhláska (CVC).

Analýza všech dat byla provedena v Praatu, kde byly detekovány jednotlivé akustické parametry. F0 a formanty byly měřeny na šesti samohláskách. V tab. 3.6 jsou uvedené výsledky hodnot frekvencí pouze pro vokál /i/. Ostatní analyzované samohlásky nelze srovnávat s českými samohláskami. Studie prokázala u žen klesání všech frekvencí: F0, F1, F2 i F3. U mužů F0 a F1 vokálu /i/ s vyšším věkem rostly, F2 a F3 klesaly.

	Vokál /i/			
	F0 [Hz]	F1[Hz]	F2[Hz]	F3[Hz]
Mladší ženy	204	373	2557	3116
Starší ženy	180	365	2360	2931
Mladší muži	114	299	2201	2986
Starší muži	142	315	2184	2804

Tab.3.6: Výsledky naměřených změn frekvencí na vokálu /i/ [9].

Počátek znělosti byl měřen na explozívech /p/, /t/, /k/, /b/, /d/, /g/. U žen VOT klesal s vyšším věkem pouze u souhlásek /b/ a /d/. U ostatních VOT rostl. U mužů VOT klesal a nebo byl konstantní s rostoucím věkem. Bylo zde ukázáno, že starší subjekty měli vyšší standardní odchylky než mladší v rámci stejného pohlaví.

Vipperla, Renals, Frankel (2010) srovnávali ve své práci [10] akustické parametry u mladších a starších mužů. Mezi analyzovanými parametry byly:

- Základní frekvence F0 [Hz],
- local jitter [%],
- local shimmer [%],
- harmonicitu.

Byly vytvořeny záznamy spontánní řeči od 23 mladších mužů ve věku 30 až 45 let a 10ti starších mužů ve věku 60 až 85 let. Z nahrávek byly vyřazeny pauzy trvající déle než 0,1 sekund. Všechny akustické parametry byly získány pomocí prostředí Praat.

Na základě ostatních studií, které potvrzovaly zlom F0 kolem 40. roku života, bylo splněné očekávání, že výsledky neprokáží výraznou změnu F0. Vhodné by bylo, kdyby v této analýze byla zařazena ještě jedna věková skupina kolem věku 20 a poté by se dal předpokládat zlom F0.

Lokální jitter vzrostl z průměrné hodnoty 1,89 % na hodnotu 2,41 %. Lokální shimmer také s rostoucím věkem vzrostl, a sice z průměrné hodnoty 10,73 % na 11,33%. Harmonicita byla naměřena u obou věkových skupin stejná.

Markó, Bóna (2010) [11] zkoumali jednak, se liší frekvenční rozsah základní frekvence mezi spontánní a čtenou podobou řeči a jak se mění rozsah s rostoucím věkem. Výzkum proběhl u 7 mladších žen, ve věku 20 až 32 let, a u 7 starších žen, ve věku 70 až 80 let. Jejich úkolem bylo přečíst 13 dlouhých vět a spontánně povídat o své práci a koníčkách. Celková doba záznamů se pohybovala kolem 15ti až 20ti minut. Dále byly nahrávky zpracovány pomocí softwaru Praat, kde byly získány minimální a maximální hodnoty základní frekvence u všech vzorků. Autoři dosáhli očekávaného výsledku, že vyšší frekvenční rozsah je obsažen ve spontánní řeči než ve čtené. Druhým výsledkem bylo porovnání poměrů maximální a minimální hodnoty F0 u mladších a starších žen. Výsledky dopadli podobně pro obě věkové skupiny. Nebyl potvrzený předpoklad, že mladší subjekty by měly mít vyšší

frekvenční rozsah. Je těžké hodnotit tento parametr v rámci příčných studií, protože se může projevat individuální intravariabilita hlasu.

Ve studii [12] je zkoumána věková závislost F0, jitteru, shimmeru a formantových frekvencí. Autoři porovnávali hodnoty F0, jitteru, shimmeru a formantů mezi dvěma věkovými skupinami. Starší skupinu tvořilo 60 subjektů (40 mužů a 20 žen) ve věku 60 až 80 let, mladší reprezentovalo 60 subjektů ve věku mezi 20 a 40 roky života. Databáze byla vytvořena ze 7500 nahraných vět, které subjekty přečetly.

Samotná analýza byla provedena v Praatu na šesti samohláskách a třech diphtongách extrahovaných z nahrávek. Výsledné trendy akustických parametrů jsou uvedeny v tab. 3.7.

	F0 [Hz]	F1 [Hz]	F2 [Hz]	F3 [Hz]	Jitter [%]	Shimmer [%]
Muži	Klesá	Klesá	Klesá	Klesá /a/ roste	Roste	Roste
Ženy	Klesá	/a/, /u/ Klesá /i/, /o/, /e/ Roste	/a/, /u/ Klesá /i/, /o/, /e/ Roste	Roste	Roste	Roste

Tab. 3.7: Výsledky trendů akustických parametrů s rostoucím věkem [12].

Báňa porovnával ve své diplomové práci [13] vybrané akustické parametry mezi šesti věkovými skupinami (věk 7 – 12, 13 – 15, 16 – 20, 21 – 30, 30 – 50, nad 50 let). Úkolem měřených subjektů bylo přečíst šest předem určených vět a nahrát s autorem jednodominový záznam spontánní řeči. Na získané databázi nahrávek byly zkoumány v Praatu charakteristiky základní frekvence F0 [Hz], akustické intenzity SPL [dB], formantových frekvencí F1 – F4 [Hz] pro dané samohlásky. Dále byla zkoumána rychlost řeči [počet slabik za sekundu]. Výzkum byl proveden pro mužské i ženské hlasy. Autor nezískal moc uspokojivé výsledky, které by se týkaly trendů. Ten se potvrdil pouze u rychlosti řeči, která v této práci s rostoucím věkem klesá u mužů i žen. Případně harmonicita vykazovala u žen s vyšším věkem menší pokles, u mužů byla stabilní. Ostatní parametry se skokově měnily a proto zde nejsou dále rozebírány.

Machač napsal velice zajímavou práci [15], zabývající se českými explozívy. Popsal zde vlastnosti explozív, jejich vznik, rozdělení, ale také popsal metodiku, podle níž poté provedl akustickou analýzu na databázi subjektů. Samotná explozíva je typická svojí délkou trvání, při zkoumání se nejčastěji analyzuje segment vokál – konsonanta – vokál, kde konsonanta představuje vybranou explozívu. Pak jsou pro výzkum důležité kromě doby trvání explozívy i doby trvání sousedních vokálů a to v souvislosti s artikulačním tempem.

Vlastní výzkum byl postaven na vytvořené databázi mluvených a čtených záznamů od dvaceti mužských i dvaceti ženských hlasů. V mluvených projevech byly použity pro jednu explozívu maximálně tři doklady. Zkoumaná explozíva byla ve spojení dvou krátkých samohlásek a byly vybírány vzorky, kde se explozíva nacházela v iniciální nebo střední poloze slova. Z výzkumu byly vyřazeny explozívy nacházející se ve finální části slov.

Dále už jsou uváděny pouze výsledky nebo metodika pro palatální explozívu *d'*, která je analyzována i v této diplomové práci. Na rozdíl od většiny dosavadních výzkumů autor zkoumá vlastnosti explozívy ve spontánní řeči, která byla získána z popisu obrazové předlohy. Byla přesně definována omezení výběru analyzovaného materiálu, aby mohly být jednotlivé zvukové záznamy srovnávány a nedocházelo ke zkreslením v konečných výsledcích.

Pro souhlásku */d'/* v pozici mezi dvěma krátkými vokály bylo zpracováno 62 zvukových záznamů (mužské i ženské hlasy). Nahrávání proběhlo ve studiu Fonetického ústavu FF UK a projevy digitalizovány vzorkovací frekvencí 22,05 kHz a samotná analýza byla provedena v programu Praat. Hranice v sekvenci vokál + explozíva + vokál byly umisťovány na základě proměnlivých fyzikálních vlastností zvukového signálu, které souvisejí s artikulačními pohyby a celkovým průběhem fonace.

V tab. 3.8 jsou uvedené výsledné průměrné hodnoty a směrodatné odchylky dob trvání explozívy a jejich fází pro souhlásku */d'/*. Kromě dob trvání jednotlivých fází závěru a exploze, které se počítají podle vzorce [15]:

$$podíl_{zaver} = \frac{t_{zaver}}{t_{exploziv}} \cdot 100 [\%] \quad (3.1)$$

	Závěr [ms]	SD	Exploze [ms]	SD	explozíva [ms]	SD	Podíl závěru [%]	Podíl exploze [%]
d' (137 dokladů)	49,1	10,6	25,4	7,9	74,6	14,0	66,0	34,0

Tab. 3.8: Doby trvání fází explozívy *d'* [15].

Pro porovnání svých výsledků uvedl autor ještě celkové průměrné délky explozívy od jiných autorů, viz tab. 3.9.

	Chlumský/ Hála	Borovičková/ Maláč	Mluvnice češtiny
d' [ms]	120	154	130

Tab. 3.9: Doby trvání explozívy *d'* od jiných autorů [15].

Mezi oběma tabulkami jsou podstatné rozdíly. Machač naměřil výrazně nižší hodnoty explozívy. To je přisuzováno odlišnému zpracovanému materiálu. Většina dosavadních výzkumů byla postavena na analýze čtených projevů řeči, které se vyznačují pomalejším tempem. Pomalejší tempo má za následek delší trvání hlásek. Výsledky Machače vycházejí jak ze vzorků mluvených projevů, tak čtených. Proto se očekává, že analýza čistě mluvených projevů bude vykazovat kratší dobu trvání explozívy.

Janda [16] provedl výzkum věkové závislosti akustických parametrů s cílem nalézt metodu, která by na základě audio nahrávek dětí byla schopná automaticky objektivně určit věk dítěte. Analyzovaná skupina byla ve věku 5 – 15 let a zkoumanými charakteristikami základní frekvence, jitter, shimmer, harmonicita, první dva formanty, rychlost řeči i podíl pauz v promluvě. Analýza byla provedena na databázi mluvených i čtených projevů. F0, jitter

a shimmer byly měřeny na prodloužené fonaci samohlásky /i/. První dvě formantové frekvence se měřily na prodloužených fonacích samohlásek /a/, /i/, /u/. Na základě těchto získaných kombinací byly vykresleny vokalické trojúhelníky pro srovnání změn mezi jednotlivými roky. Pro porovnání rychlostí řeči byla nahrána dětská říkanka, kterou všechny děti dobře znaly. Podíl pauz v promluvě byl naopak získán ze spontánního vyprávění o daných obrázcích. Procentuální zastoupení pauz v promluvě bylo vypočteno podle vzorce:

$$pp = \frac{l_t - l_a}{l_t} \cdot 100 [\%] \quad [16], \quad (3.2)$$

kde l_t je celková délka promluvy, l_a je suma délek všech artikulovaných úseků promluvy.

V tab. 3.10 jsou uvedené výsledné trendy jednotlivých parametrů vycházející z Jandovy studie. Pauzy a rychlost řeči se zde předpokládají opačné než u analýzy dospělých jedinců, zde mají starší děti lepší vyjadřovací schopnosti a proto výsledky odpovídají předpokladům.

Trend s vyšším věkem	F0 [Hz]	Jitter [%]	Shimmer [%]	F1 [Hz]	F2 [Hz]	Rychlost řeči [Slabiky/s]	Pauzy [%]
Dívky	Klesá	Roste mezi 13.-15. Rokem	Roste prudce mezi 13.-15. Rokem	Klesá pro /a/, /i/, /u/	Klesá pro /a/, Roste pro /i/, /u/	Roste	Klesá
Chlapci	Klesá	Roste prudce mezi 13.-15. Rokem	Roste prudce mezi 13.-15. Rokem	Klesá pro /a/, /i/, /u/	Klesá pro /a/, Roste pro /i/, /u/	Roste	Klesá

Tab. 3.10: Trendy akustických parametru Jandovy studie [15].

3.2.2 Přehled výsledků příčných studií

V tab. 3.2.1 jsou uvedené výsledky trendů pro jednotlivé akustické parametry, které byly analyzovány v příčných studiích zmíněných výše v přehledu literatury.

Tab. 3.2.1: Shrnutí výsledků příčných studií.

Základní frekvence F0 [Hz]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	F0 [Hz]	SD [Hz]
[6]	Starší muži (75,4) Normované hodnoty mladších	-	Prodloužená fonace samohlásky /a/	Klesá 145-128	Roste 23-29

Tab. 3.2.1: Pokračování

Základní frekvence F0 [Hz]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	F0 [Hz]	SD [Hz]
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov	20-40 Klesá 50-80 Roste	
[9]	12 mladších (25,5 let) 27 starších (75,2 let)	50 let	Čtení 22 segmentů CVC, 6 samohlásek /i/	Roste 114-142	
[10]	23 mladších (30-45 let) 10 starších (60-85 let)	30-40 let	Spontánní řeč	Stabilní	
[12]	40 mladších (20-40) 40 starších (60-80)	30-40 let	Čtené projevy, 6 samohlásek, 3 diphtongy	Klesá	
[16]	Chlapci (5-15 let)	1 rok	Mluvené i čtené projevy Prodloužená fonace /i/	Klesá	
ŽENSKÉ HLASY					
[6]	Starší ženy (74,8) Normované hodnoty mladších	-	Prodloužená fonace samohlásky /a/	Klesá 244-188	Roste 27-42
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov	20-50 Stabilní 60-80 Roste 80- Klesá	
[9]	15 mladších (25,5 let) 32 starších (75,2 let)	50 let	Čtení 22 segmentů CVC, 6 samohlásek /i/	Klesá 204-180	
[12]	20 mladších (20-40 let) 20 starších (60-80)	30-40 let	Čtené projevy, 6 samohlásek, 3 diphtongy	Klesá	
[16]	Dívky (5-15 let)	1 rok	Mluvené i čtené projevy Prodloužená fonace /i/	Klesá	

Tab. 3.2.1: Pokračování

Jitter, shimmer [%]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	Jitter [%]	Shimmer [%]
[6]	Starší muži (75,4) Normované hodnoty mladších	-	Prodloužená fonace samohlásky /a/	Roste 0,6-2,1	Roste 2,5-5,5
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov		Do 40ti Roste 40-70 Klesá Stabilní Po 80 Klesá
[10]	23 mladších (30-45 let) 10 starších (60-85 let)	30-40 let	Spontánní řeč	Roste 1,9-2,4	Roste 10,7-11,3
[12]	40 mladších (20-40) 40 starších (60-80)	30-40 let	Čtené projevy, 6 samohlásek, 3 diphtongy	Roste	Roste
[16]	Chlapci (5-15 let)	1 rok	Mluvené i čtené projevy Prodloužená fonace /i/	13-15 prudce roste	13-15 prudce roste
ŽENSKÉ HLASY					
[6]	Starší ženy (74,8) Normované hodnoty mladších	-	Prodloužená fonace samohlásky /a/	Roste 0,6-2,0	Roste 2,0-5,3
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov		Do 50 stabilní 50-70 Klesá Po 70 Roste
[12]	20 mladších (20-40 let) 20 starších (60-80)	30-40 let	Čtené projevy, 6 samohlásek, 3 diphtongy	Roste	Roste
[16]	Dívky (5-15 let)	1 rok	Mluvené i čtené projevy Prodloužená fonace /i/	13-15 Roste	13-15 prudce roste

Tab. 3.2.1: Pokračování

Formanty F1, F2 [Hz]					
MUŽSKÉ HLASY					
Studie	Měřené subjekty	Časový rozdíl	Metodika	F1 [%]	F2 [Hz]
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov	Klesá	Roste
[9]	12 mladších (25,5 let) 27 starších (75,2 let)	50 let	samohláska /i/	Roste 299-315	Klesá 2201-2184
[12]	40 mladších (20-40) 40 starších (60-80)	30-40 let	Čtené projevy /a/, /e/, /i/, /o/, /u/	Klesá	Klesá
[16]	Chlapci (5-15 let)	1 rok	Prodloužená fonace /a/, /i/, /u/	Klesá	/a/ klesá /i/, /u/ Roste
ŽENSKÉ HLASY					
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov	Klesá	Roste
[9]	15 mladších (25,5 let) 32 starších (75,2 let)	50 let	samohláska /i/	Klesá 373-365	Klesá 2557-2360
[12]	20 mladších (20-40 let) 20 starších (60-80)	30-40 let	Čtené projevy /a/, /e/, /i/, /o/, /u/	/a/, /u/ Klesá /e/, /i/, /o/ roste	/a/, /u/ Klesá /e/, /i/, /o/ roste
[16]	Dívky (5-15 let)	1 rok	Prodloužená fonace /a/, /i/, /u/	Klesá	/a/ klesá /i/, /u/ Roste
Počátek znělosti VOT [ms]					
Studie	Měřené subjekty	Časový rozdíl	Metodika	VOT [ms]	Hodnota VOT [ms]
[15]	Ženy, muži	-	Spontánní řeč + čtená VCV, explozíva d' (čtený projev – jiné studie)		74,6 120 154 130
MUŽSKÉ HLASY					
[9]	12 mladších (25,5 let) 27 starších (75,2 let)	50 let	12 segmentů CVC /p/, /t/, /k/, /b/, /d/, /g/	Klesá, stabilní	
ŽENSKÉ HLASY					
[9]	15 mladších (25,5 let) 32 starších (75,2 let)	50 let	12 segmentů CVC /p/, /t/, /k/, /b/, /d/, /g/	/b/, /d/ Klesá /p/, /t/, /k/, /g/ Roste	

Tab. 3.2.1: Pokračování

Rychlost řeči [slabik/s], pauzy v promluvě [%]					
Studie	Měřené subjekty	Časový rozdíl	Metodika	Rychlost řeči [slabik/s]	Pauzy [%]
[8]	Muži, ženy	-	explozíva	Roste	
MUŽSKÉ HLASY					
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov	Prodlužují se segmenty	
[13]	Muži 16-50+	10 – 20 let	1 min spontánní řeči, 6 čtený vět	Klesá	
[16]	Chlapci (5-15 let)	1 rok	Dětská říkanka rychlost Popis obrázku pauzy	Roste	Klesá
ŽENSKÉ HLASY					
[7]	8 věkových skupin (20-90 let)	10 let	6 izolovaných slov	Prodlužují se segmenty	
[13]	Ženy 16-50+	10 – 20 let	1 min spontánní řeči, 6 čtený vět	Klesá	
[16]	Dívky (5-15 let)	1 rok	Dětská říkanka rychlost Popis obrázku pauzy	Roste	Klesá

Příčných studií co do počtu je více než studií podélných. Nutností je ale zastoupit v každé zkoumané věkové skupině dostatečný počet subjektů, aby se potlačila individuální intravariabilita hlasů. Tabulka 3.2.1 ukazuje i větší počet zkoumaných akustických parametrů.

Na základě rešerše zkoumané problematiky z podélných a příčných studií byly navrženy následující akustické parametry, které jsou experimentálně ověřeny na vytvořené databázi ženského hlasu v rámci jeho podélné studie:

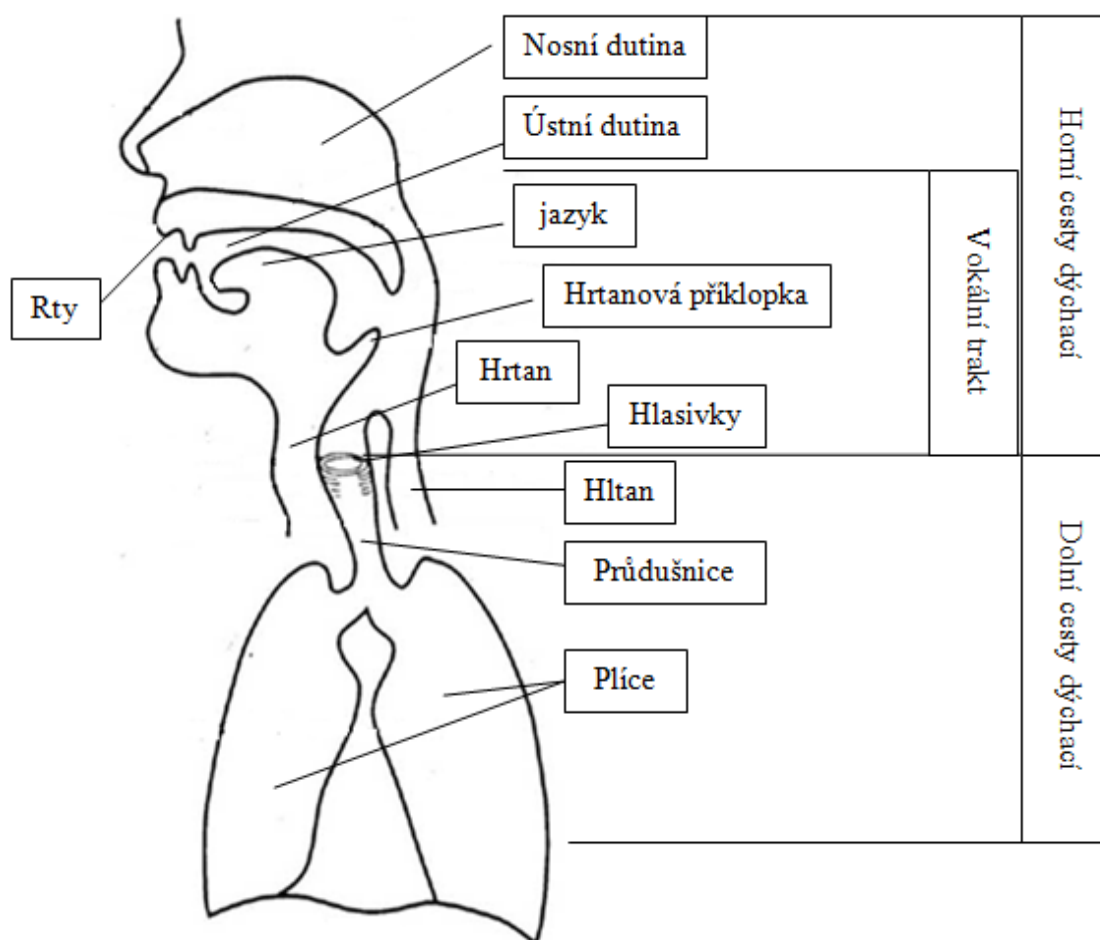
- Základní frekvence [Hz] z celých promluv z prodloužené fonace /a/.
- Jitter, shimmer [%] z prodloužené fonace samohlásky /a/.
- F1, F2 z prodloužených fonací samohlásek /a/, /e/, /i/, /o/, /u/.
- Vokální trojúhelník vytvořený z F1 a F2 všech vokálů.
- Počátek znělosti VOT [ms] na segmentu vokál – explozíva – vokál
- Rychlost řeči [slabik/s] z delších promluv

Předpokládané trendy navržených parametrů jsou následující:

- F0 bude s rostoucím věkem klesat.
- F1 u všech vokálů bude klesat, F2 bude klesat u /a/, růst u /i/, /u/
- VOT explozív se bude zkracovat, délka vokálů by se měla prodlužovat, rychlost celkově s vyšším věkem prodlužovat

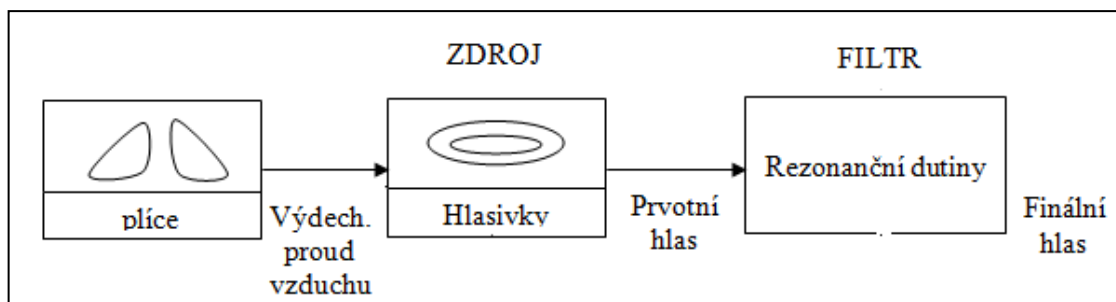
4. Lidský hlas

Zvuk je vlnění vytvořené kmitáním tělesa. V případě hlasu jde o kmitající se hlasivky. Nejde o aktivní orgán, ale o pasivní kmitání díky vzduchu proudícímu z plic. Hlasivky mají schopnost zcela otevřít, částečně uzavřít nebo zcela uzavřít cestu unikajícímu vzduchu z plic. Při částečném uzavření hlasivky při výdechu kmitají a vytvářejí tak základní hlasivkový tón [21]. Artikulační orgány jsou znázorněny na obrázku 4.1.



Obr.4.1: Artikulační orgány.

Hlasivky jsou zdrojem vzniku primárního hlasu, rezonanční dutiny naopak slouží jako filtr, který vlivem rezonancí v dutinách vokálního traktu upravuje základní hlas ve finální. Frekvence vibrací hlasivek závisí na vzduchu vycházejícího z plic, hmotnosti a pružnosti hlasivek, ale také na velikosti plochy hlasivkové štěrbině ve volném stavu. Vysvětlení vzniku hlasu pomocí teorie zdroje a filtru je znázorněno na obr. 4.2.



Obr. 4.2: Schéma vzniku hlasu podle teorie zdroje a filtru.

Ze zdroje získáme základní frekvenci, neboli výšku tónu, základní intenzitu, délku fonace, primární kvalitu hlasu (dyšnost, chraplavost, hrubost), primární barvu hlasu. Po průchodu rezonančními dutinami už dokážeme rozlišit samohlásky, souhlásky. Máme konečnou intenzitu, konečnou barvu, známe formantové frekvence.

4.1 Artikulace

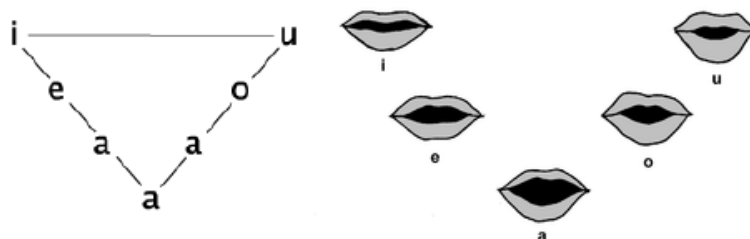
Proces artikulace probíhá v nadhrtanových dutinách, do kterého se zapojují artikulační orgány. Mezi ně se řadí rty, zuby, tvrdé i měkké patro, zadní stěna hrdelní dutiny, spodní čelist a jazyk. Podle činnosti se dělí na vokalickou a konsonantickou.

Pokud získáme zvuk s převahou tónové složky, pak jde o artikulaci vokalickou. Jde pouze o změnu velikosti a tvaru dutiny ústní, kde nedochází k tvorbě výrazných překážek.

U konsonantické artikulace dochází k tvorbě překážek unikajícímu proudu vzduchu. Výsledkem je zvuk s převahou šumové složky [18].

4.1.1 Samohlásky (vokály)

Nejvýznamnější artikulační charakteristikou českých samohlásek je poloha a pohyb jazyka [18]. Posunem dochází ke změně poměrů objemů dutiny ústní a hrdelní. Postavení jazyka v ústech zobrazuje obecné schéma vokálů ve formě vokalického trojúhelníku. Vychází z vertikálního a horizontálního posunu jazyka z klidové polohy při tvoření jednotlivých samohlásek. Na obr. 4.3 je znázorněné Hellwagovo schéma samohlásek a postavení úst pro české samohlásky /a/, /e/, /i/, /o/, /u/.



Obr. 4.3: Vlevo: Hellwagův trojúhelník, vpravo: práce úst při artikulaci českých samohlásek [19].

Z akustického hlediska mají samohlásky tónovou strukturu. Jsou všechny znělé. Jde o složený tón vzniklý z laryngálního hlasu (základního tónu) a rezonancí (formantů) vzniklých v nadrtanových dutinách.

Nejvýznamnější akustickou charakteristikou samohlásek jsou formanty. První dva formanty, které rozlišují jednotlivé české samohlásky. Každý vokál je určen jiným poměrem prvních dvou formantů (tónových složek). Frekvence a intenzita formantů závisí na uspořádání, délce, tvaru a průřezu dutiny ústní i hrdelní. Hlavní anatomické části, které způsobují změny jejich frekvence a intenzity jsou rty, čelisti, jazyk a měkké patro (uzavírá vstup vzduchového proudu do dutiny nosní pro všechny české samohlásky). Pomocí těchto artikulátorů může řečník měnit velikost nadrtanového prostoru a tím modifikovat formantovou strukturu vytvářených zvuků (vytvářet různé samohlásky) [18].

Spektrální vlastnosti jednotlivých českých samohlásek jsou uvedeny níže [18] a hodnoty prvních dvou formantů v tab. 4.1.

- **Střední nízké samohlásky** /a/, /á/ jsou typické minimální vzdáleností prvních dvou formantů F_1 , F_2 ($F_2 / F_1 \doteq 1,6$). Změny formantové struktury mezi krátkou a prodlouženou fonací samohlásky jsou z percepčního hlediska zanedbatelné.
- **Přední středové samohlásky** /e/, /é/ mají velký rozdíl mezi oběma formanty ($F_2 / F_1 \doteq 3,2$).
- **Přední vysoké samohlásky** /i/, /í/ mají největší vzdálenost hodnot prvních dvou formantů ($F_2 / F_1 \doteq 6,5$).
- **Zadní středové samohlásky** /o/, /ó/ mají oba formanty docela blízko ($F_2 / F_1 \doteq 1,8$).
- **Zadní vysoké samohlásky** /u/, /ú/ jsou typické velmi malou vzdáleností mezi oběma formanty a ty se pohybují na nízkých frekvenčních hodnotách.

Vokál	F_1 [Hz]	F_2 [Hz]
/i/, /í/	300 – 500	2000 – 2800
/e/, /é/	480 – 700	1560 – 2100
/a/, /á/	700 – 1100	1100 – 1500
/o/, /ó/	500 – 700	850 – 1200
/u/, /ú/	300 – 500	600 – 1000

Tab.4.1: Hodnoty F_1 a F_2 pro české samohlásky [17].

Uvedené hodnoty formantů jsou výsledkem měření širokého vzorku populace. Pro různé řečníky se hodnoty mohou obecně dost odlišovat, protože závisejí na tvaru, velikosti a délce celého hlasového traktu člověka. Jeho rozměry se pro různé lidi velmi odlišují. V rámci jednoho řečníka ale k takovýmto překryvům nedochází.

Typickou vlastností samohlásek je delší doba trvání v porovnání se souhláskami. Trvání samohlásek ale kolísá vlivem realizace suprasegmentálních rysů a podle aktuálního tempa mluvení. Amplituda (intenzita) samohlásek je výrazně vyšší než je u souhlásek. Na hodnotě intenzity se podílí emoce řečníka, druh promluvy i okolní prostředí [18].

4.1.2 Souhlásky (konsonanty)

Akustické signály souhlásek jsou příznačné přítomností charakteristického šumu v porovnání s akustickými signály samohlásek i menší amplitudou. Při artikulaci dochází ke ztěžování průchodu výdechovému proudu vzduchu hlasovým ústrojím buďto vytvořením překážky nebo zúžením hlasového traktu v určitém místě hlasového ústrojí. Podle místa vytvoření překážky nebo zúžení vznikají různé zvuky, jejichž podstatou je šum [18].

České souhlásky se dělí podle čtyř základních charakteristik:

- Způsob artikulace
 - Závěrové (explozivní): /p/, /t/, /tʰ/, /k/, /b/, /d/, /dʰ/, /g/, /m/, /n/, /ŋ/.
 - Úžinové (konstruktivní): /f/, /v/, /s/, /š/, /z/, /ž/, /j/, /ch/, /h/, /l/, /r/, /ř/.
 - Polozávěrové (semiokluzivní): /c/, /č/, /dz/, /dž/.
- Místo artikulace
 - Retné (labiální): /p/, /b/, /m/, /f/, /v/.
 - Dásňové (alveolární): /t/, /d/, /n/, /s/, /z/, /š/, /ž/, /c/, /dzž/, /č/, /l/, /r/, /ř/.
 - Patrové: /tʰ/, /dʰ/, /ŋ/, /j/, /k/, /g/, /ch/.
 - Tvrđopatrové (palatální): /tʰ/, /dʰ/, /ŋ/, /j/.
 - Hrtanové (glotální): /h/.
- Postavení měkkého patra
- Činnost hlasivek

Při vytváření většiny českých souhlásek prochází výdechový proud vzduchu dutinou hrdelní a ústní. Pokud měkké patro uvolní průchod vzduchu do dutiny nosní, pak se účastní artikulace i tato dutina a vznikají nazály [18].

V rámci analýzy, která je provedena v této práci je zkoumán akustický parametr počátek znělosti. Je měřen ze sekvence vokál – exploziv – vokál, proto jsou dále popsány jen explozivní.

4.1.2.1 Explozivní

Explozivní jsou přechodové zvuky, jejichž artikulace se skládá ze dvou fází: okluze (závěr, překážka) a exploze (uvolnění překážky). Během okluze je vokální trakt plně uzavřen, takže se před překážkou hromadí výdechový proud vzduchu a stoupá intraorální tlak. Vzhledem k předcházející hlásce v segmentu VEV (vokál – exploziv – vokál) dochází ke zřetelnému oslabení amplitudy [15]. Při explozi dojde k uvolnění překážky a rychle unikající proud vzduchu mezi oddalujícími se artikulátory způsobí vyrovnání intraorálního a atmosférického tlaku a nahromaděný vzduch pod závěrem po určitou dobu prochází úzkým prostorem mezi artikulačními orgány, čímž vzniká šum [20].

Temporální vlastnosti exploziv bývají zkoumány v souvislosti s místem artikulace a znělostí. Znělé a neznělé exploziv se odlišují jednak znělostí/ neznělostí a jednak stavbou a pohyblivostí artikulačních orgánů a podobou cílové artikulační pozice [15].

Poměr trvání okluze a exploze ovlivňuje míra pohyblivosti artikulujících orgánů při zaujímání a opouštění cílové artikulační pozice a podoba cílové artikulační pozice (velikost a

místo dotykové plochy). Rozdílná znělost je určena způsobem koordinace laryngálních a orálních úkonů, vzájemnými poměry subglottálního, intraorálního a atmosférického tlaku, rychlostí oddalování artikulačních orgánů při uvolňování závěru nebo změnou rychlosti proudění vzduchu [15].

Interpretace temporálních vlastností explozív níže je převzata z práce [15, 20] a opírá se o výzkumy dalších autorů. Shrnutí rozdílů mezi znělými a neznělými explozívy je v tab. 4.2.

- **U neznělých explozív** jsou v podstatné části závěrové fáze hlasivky oddáleny, a tím pádem není rozdíl mezi subglottálním a supraglottálním tlakem. V průběhu závěrové fáze se začíná hromadit výdechový vzduch pod závěrem a dochází k narůstání intraorálního tlaku až dojde k dosažení tlakového maxima v nadhrtanových dutinách. Vlivem napětí artikulačního svalstva zůstává objem nadhrtanových dutin a tlak vzduchu v nich po dobu závěru téměř konstantní.
 - Po celou dobu trvání se ve spektru nevyskytuje základní frekvence ani formantová struktura. Mají výrazný explozivní šum.
- **U znělých** se mechanika pohybů artikulačního ústrojí snaží udržet fonaci po celou dobu závěru. Pro udržení kmitání hlasivek výdechovým proudem vzduchu je nutné udržet dostatečný rozdíl subglottálního a supraglottálního tlaku. Proto dochází k částečnému zvětšování objemu nadhrtanových dutin a tím pádem je menší napětí v artikulačním svalstvu. Jsou daná jistá omezení ve zvětšování objemu vokálního traktu, proto je i trvání explozív omezené.
 - Po celou dobu jejich trvání je přítomna základní frekvence bez formantové struktury. Šum v explozivní fázi bývá slabý nebo úplně chybí.

	Neznělá explozíva	Znělá explozíva
Hlasivky	Oddáleny, volný průchod vzduchu	Semknuty, pozvolný průchod vzduchu
Prostor mezi hlasivkami a závěrem při okluzi	Konstantní velikost	Mírně se zvětšuje
Rozdíl subglottálního a supraglottálního vzduchu	Po vytvoření závěru se vyrovná	Zůstává po celou dobu závěru
Intraorální tlak	Vyšší	Nižší
Dotyková plocha	Větší	Menší
Artikulační svalstvo	Napjatější	Povolenější
Trvání závěru	Delší	Kratší
Uvolnění závěru	Pomalejší	Rychlejší
Vyrovnání intraorálního a atmosférického tlaku	Trvá delší dobu	Kratší dobu
Trvání exploze	Delší	Kratší

Tab. 4.2: Rozdíly při tvorbě českých znělých a neznělých explozív [15].

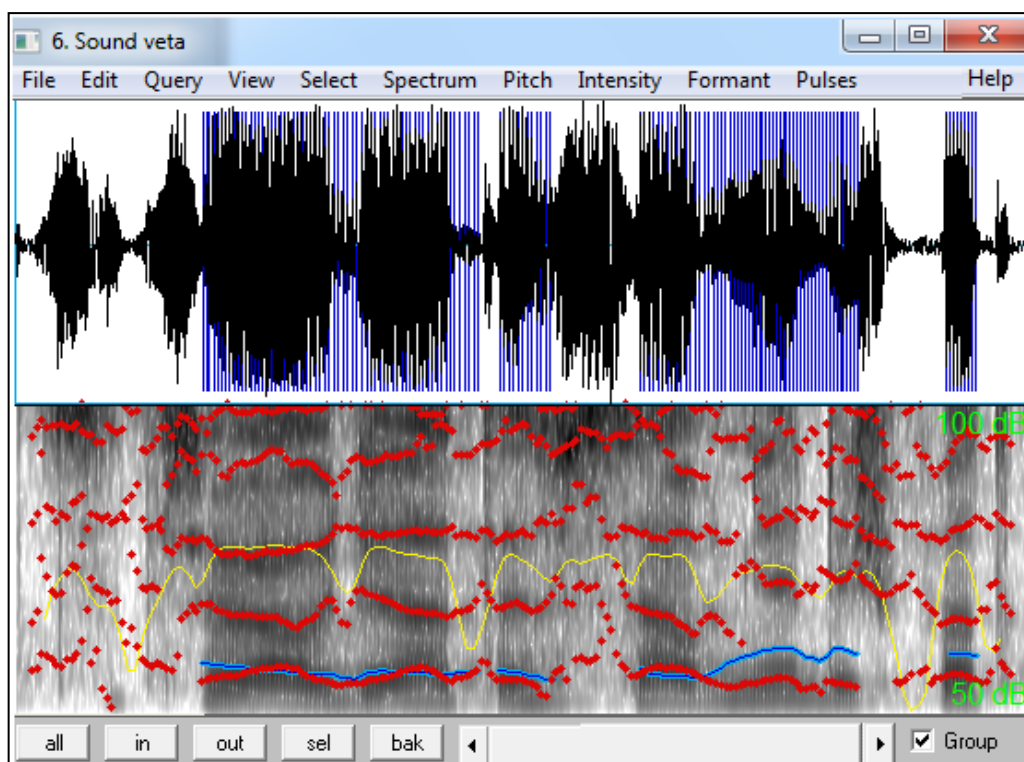
5. Praat

Praat je volně dostupný software z oficiálních internetových stránek www.praat.org, který je určený pro zpracování řeči. „Praat“ v holandštině znamená „mluvit“. Byl vytvořen a je dále vyvíjen prof. P. Boersmou a dr. D. Weeninkem z Univerzity v Amsterdamu. Z oficiálních stránek lze stáhnout také zdrojový kód balíku Praat a přidávat vlastní nové funkce. Kód je napsaný v jazyce C++. K softwaru je dostupná dobře zpracovaná nápověda pro ovládání programu v anglickém jazyce [22].

5.1 Možnosti Praatu

Praat obsahuje rozsáhlou řadu funkcí pro analýzu a rekonstrukci řečového signálu. Lze zpracovávat soubory ve formátu .wav. Využívá základní výpočetní algoritmy pro aplikaci dat na signál. Nabízí i možnost psát skripty pro zpracování většího počtu signálů.

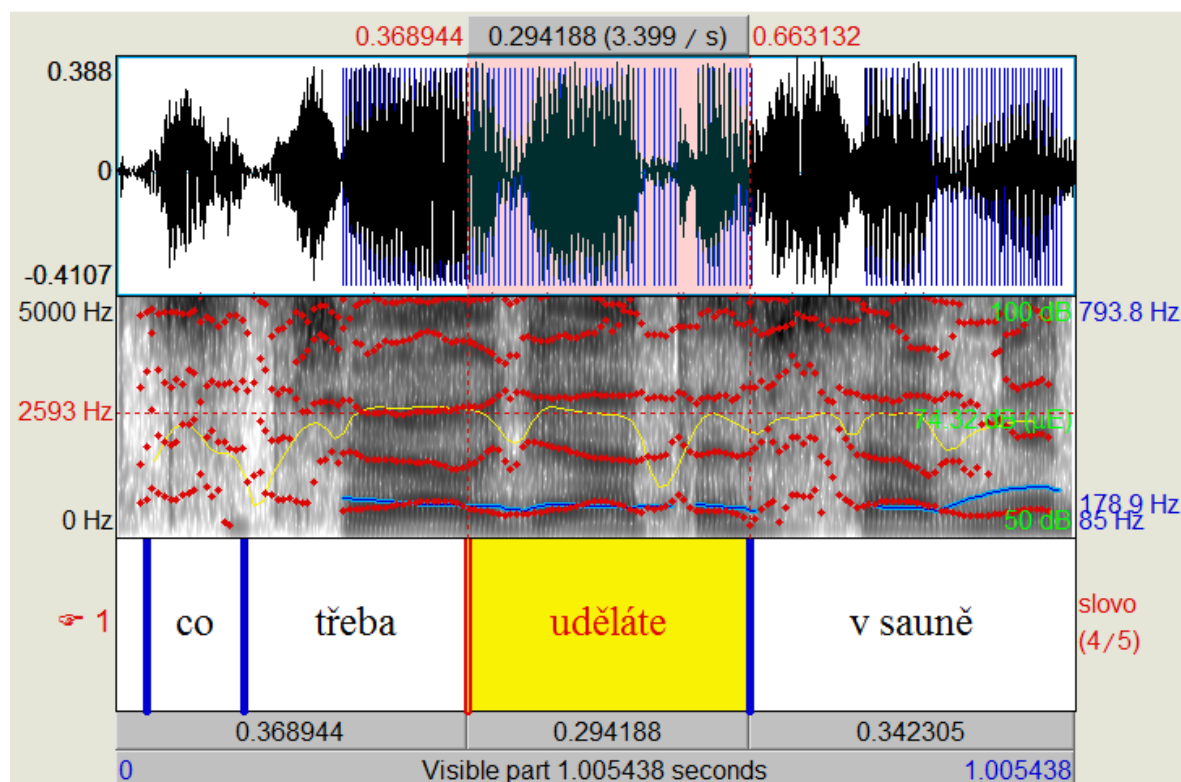
- **Praat Objects window** je startovním místem pro zpracování zvukového souboru. Pomocí tohoto menu lze otevírat, vytvářet a ukládat soubory.
- **Editor window** je otevřeno v menu po označení vybraného wav souboru výběrem „View & Edit“. V tomto okně uživatel provádí většinu akustických analýz. Okno je rozděleno na polovinu, kde v horní je vykreslen řečový signál v čase a ve spodním jeho spektrum, viz obr. 5.1.



Obr. 5.1: Editor okno pro práci se signálem.

- **Základní frekvence** ve spektru na obr. 5.1 odpovídá modrému průběhu. Hodnoty se zjistí výběrem „Pitch“ v hlavičce editorového okna. Dále se zde dají nastavovat meze pro detekování F0, získat průměrná minimální i maximální hodnota F0 po označení vybraného úseku zvukového signálu, dá se vykreslit F0 samostatně do obrázku v dalším okně Praat picture.
- **Formantovým frekvencím** odpovídají červené neaproximované křivky. Standardní nastavení v Praatu je pro detekci pěti formantů (F1 - F5) do frekvence 5500 Hz. Nastavení počtu formantů a frekvenčního omezení se dá upravovat v záložce „Formant – Formant settings..“.
- **Intenzita** je zobrazena jako žlutá křivka. Průběh hodnot, maximální i minimální průměrné hodnoty se dají zjistit výběrem „Intensity“.
- Záložka „Pulses“ nabízí výsledky analýzy označeného signálu jako „Voice report“. V této zprávě jsou obsaženy údaje o základní frekvenci (medián, průměr, standardní odchylka, minimální F0 a maximální F0), počet pulzů označeného úseku signálu (průměrná perioda, standardní odchylka), poměr znělých a neznělých částí v signálu, jitter, shimmer a harmonicitu.
- Záložky na spodní liště (obr. 5.1 dole) „sel“ a „bak“ se používají pro přiblížení, zvětšení vybraného úseku signálu a vrácení zpět k původnímu signálu. Např. když je třeba stanovit hranice počátku znělosti, musí se signál přiblížit.

Menu „Praat Objects“ nabízí možnost vytvořit textový soubor označením wav souboru a výběrem „Annotate – To Textgrid..“. Tím se vytvoří nový soubor, označením obou souborů a „View & Edit“ je zobrazen signál se spektrem i s textovým popisem, obr. 5.2.



Obr. 5.2: Editor okno s vykresleným zvukovým signálem i textovým popisem.

6. Věkově závislé parametry

Metodologie akustických analýz je postavena na zkoumání tří základních vlastností řečových signálů. Na fonačních, které jsou spojeny s tvorbou prvotního signálu hlasu, na artikulačních, které souvisí s tvorbou artikulace, rozlišováním samohlásek a souhlásek a na prozodických vlastnostech. Ty nepopisují význam řeči, co řečník říká, ale jak to říká. V následujících podkapitolách jsou uvedené a popsány vybrané akustické parametry na základě rešerše v kapitole 3. Dále je ověřena experimentálně jejich věková závislost na databázi ženského hlasu v kapitole 7. Vybrané akustické parametry jsou základní frekvence (F0), frekvenční a amplitudová variabilita základní frekvence (jitter, shimmer), první dvě formantové frekvence (F1, F2), velikost vokálního prostoru (VSA), artikulační index (VAI), počátek znělosti (VOT), rychlost řeči a podíl pauz v řečovém signálu.

6.1 Fonační vlastnosti řeči

Základní vlastnosti lidského hlasu jsou dány fyziologií hlasového ústrojí. Fonační (hlasové) ústrojí je uloženo v hrtanu a jeho funkcí je vytvářet základní hlas. Zdrojem pro tvorbu hlasu jsou hlasivky. Významnou vlastností hlasivek je jejich délka. Čím jsou kratší, tím rychleji kmitají a běžný mluvní hlas je vyšší. Vlivem hormonálních změn v organismu se mění základní poloha hlasu dospívání (mutace) a další změny nastávají ve stáří. Změny síly hlasu vznikají zesílením výdechového proudu, čímž dojde ke zvětšení amplitudy kmitajících hlasivek. Schopnost silové i tónové modulace je částečně vrozená, ale také se dá dosáhnout pravidelným cvičením (herci, zpěváci) [19].

Standardní fonační analýza zahrnuje měření rozsahu základní frekvence, průměrných hodnot F0 a standardních odchylek. S průběhem F0 souvisejí i jeho frekvenční a amplitudové odchylky, proto jsou měřeny parametry jitter, shimmer a harmonický šum (NHR).

6.1.1 Základní frekvence F0

Základní frekvence vzniká kmitáním hlasivek. Je pozorována ve znělých úsecích řeči. Změny v základní frekvenci jsou způsobeny změnami v rychlosti kmitání hlasivek. Z časového průběhu zvukového signálu i ze spektra je nejdříve určena základní perioda, která je ovlivněna vlastnostmi hlasivek (průměrem, délkou, hmotností a pružností) a na základě vzorce vypočtena základní frekvence.

$$F0 = \frac{1}{T0} [\text{Hz}] \quad (6.1)$$

Rozsah F0 řeči je rozdílný pro muže, ženy i děti a je uveden v tab. 6.1.

	Muži	Ženy	Děti
F0[Hz]	80 - 200	150 - 350	200 - 500

Tab. 6.1: Rozsah F0 v řeči [17].

Základní frekvence je nejčastěji analyzovaným parametrem v akustických studiích. Podélné a příčné se ve výsledcích průběhu F0 s rostoucím věkem liší. Podélné studie mužských i ženských hlasů tvrdí, že F0 s rostoucím věkem klesá [1, 2, 3, 4, 5]. Výsledky příčných studií mužského hlasu tvrdí klesání s věkem [7, 12], růst s věkem [9] nebo stálost základní frekvence v průběhu let [10]. U ženského hlasu většina příčných studií potvrzuje klesající trend s věkem [6, 9, 12]. Ve studii [7] je F0 do věku 50 stabilní, 60 – 80 let roste a po 80. roce klesá.

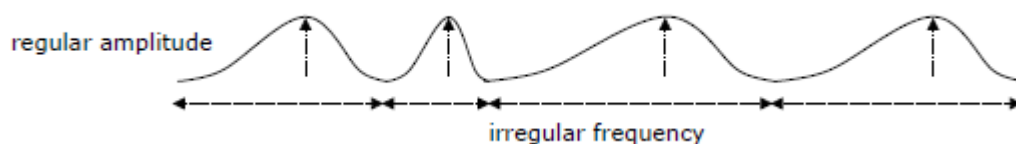
Rozdílnost výsledků je dána různou metodikou měření parametru F0. Je hodnocena z celých promluv (spontánní, čtená) nebo na prodloužených fonacích samohlásek /a/, /i/.

6.1.2 Frekvenční variabilita jitter

Kvalita hlasu (fonace) se vyhodnocuje pomocí parametru jitter. Obvykle se stanovuje z prodloužené fonace hlásky /a/ nebo /i/. Jitter popisuje stupeň poruchy periodicity základní frekvence, způsobený změnami v rychlosti kmitání hlasivek. Nepravidelné změny frekvence jsou znázorněny na obr. 6.1. Jitter lze definovat jako krátkodobou nepravidelnost délek jednotlivých period řečového signálu. Lokální kolísání periody je určeno podílem průměrné absolutní hodnoty rozdílu mezi sousedními periodami a průměrné délky periody podle následujícího vzorce 6.2.

$$J = \frac{\frac{1}{N-1} \left[\sum_{i=1}^{N-1} |T_i - T_{i+1}| \right]}{\frac{1}{N} \sum_{i=1}^N T_i} \cdot 100 \text{ [Hz]}, \quad (6.2)$$

kde N je počet analyzovaných period, T je doba trvání periody [16].



Obr. 6.1: Frekvenční variabilita [8].

Trendy tohoto parametru s věkem u zdravých jedinců jsou důležité pro včasnou diagnózu karcinomu hrtanu a celkových poruch hlasivek.

Většina studií věkové závislosti jitteru potvrdila rostoucí trend s věkem u mužů i žen [6, 10, 12]. Studie [16] hodnotila parametry u dívek a chlapců do 15ti let. Růst jitteru byl potvrzen ve věku 13 až 15 let. U chlapců byl prudší růst v důsledku mutace. Celkově se hodnoty jitteru zdají být závislé spíše na zdravotním a fyziologickém stavu mluvčích než na věku [8].

V této podélné studii ženského hlasu je předpokládán pozvolný nárůst frekvenčního kolísání s rostoucím věkem v důsledku únavy hlasu a nestálostmi hlasivek.

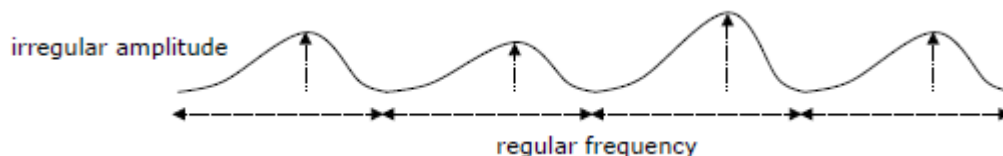
6.1.3 Amplitudová variabilita shimmer

Shimmer označuje odchylky ve velikosti maximálních amplitud v jednotlivých periodách. Určuje míru rozdílnosti energie v jednotlivých periodách signálu. Amplitudové odchylky jsou vykresleny na obr. 6.2. Lokální kolísání amplitudy je určeno podílem průměrného rozdílu absolutních hodnot amplitud po sobě jdoucích period a průměrné amplitudy podle vzorce 6.3. Odchylky jsou měřeny nejčastěji společně s jitterem na prodloužených fonacích samohlásek /a/ nebo /i/.

Amplituda se mění v závislosti i na emocích mluvčího. U herců, kteří pracují s hlasem každý den, se amplituda základního hlasu mění v závislosti na dané roli [19] .

$$S = \frac{\frac{1}{N-1} \left[\sum_{i=1}^{N-1} |A_{i+1} - A_i| \right]}{\frac{1}{N} \sum_{i=1}^N A_i} \cdot 100 \text{ [Hz]} , \quad (6.3)$$

kde N je počet analyzovaných period, A je největší rozdíl mezi maximem a minimem amplitudy v periodě.



Obr. 6.2: Amplitudová variabilita [8].

Většina studií věkové závislosti tohoto parametru v kapitole 3 uvádí růst amplitudového kolísání s rostoucím věkem [6, 10, 12, 16]. Ve výsledcích se lišila akorát studie [7], kde byl shimmer u mužů stabilní do čtyřiceti let, poté klesal, u žen byl stabilní do padesáti let, pak klesal do věku 70 a po 70. roce narůstal.

Změny související s věkem u zdravých jedinců jsou důležité pro včasnou diagnózu lézí na hlasivkách, karcinomů. S rostoucím věkem je předpokládán nárůst amplitudového kolísání u zdravých jedinců související s únavou hlasu a nestálostmi hlasivek.

6.2 Artikulační vlastnosti řeči

Artikulační vlastnosti jsou spojené s tvorbou hlásek. Hlas vycházející z hlasivek nemá barvu lidského hlasu. Charakteristická znění, individuální pro jednotlivce, jsou získána až průchodem nadhrtanovými dutinami, které mohou měnit v určitých mezích svůj tvar a objem. Do nich vstupují impulsy a šum nesené výdechovým proudem vzduchu z plic. Jednotlivým akustickým rezonančním frekvencím těchto dutin odpovídají formanty [17]. Samohlásky jsou diferenciovány na základě kombinace prvních dvou formantů (F1, F2), proto je v akustických analýzách hodnocena plocha vokálního prostoru určená vzdálenostmi mezi jednotlivými

samohláskami ve vokalickém trojúhelníku (VSA). Dalším zkoumaným parametrem je artikulační index, který určuje míru rozlišitelnosti jednotlivých samohlásek (VAI).

Formantové frekvence

Formanty odpovídají akustickým rezonančním frekvencím dutin, kterými prochází základní hlasivkový tón z hlasivek společně s šumem. Ve spektrogramech se zobrazují jako tmavé pásy. Pořadí formantů se určuje od nejnižší frekvence. Přičemž se musí rozlišovat mezi frekvencí základního tónu a prvním formantem [17].

Podélné studie se shodly ve výsledcích průběhů (F1, F2) u mužů [5] i žen [2, 3, 5]. Jejich trend s rostoucím věkem je klesající. Studie [5] dokonce tvrdí, že průběh F1 aproximuje průběh F0, který byl také klesající. Formanty byly měřeny v záznamech spontánní řeči i čtené. Hodnoceny byly znělé segmenty vět, obsahující vokály /a/, /i/, /u/.

Příčné studie už se mírně rozcházejí ve výsledcích. U mužských hlasů F1 klesala s rostoucím věkem [7, 12, 16] a rostla [9]. F2 klesala s rostoucím věkem [9, 12] a rostla [7]. U ženských hlasů všechny studie potvrdily klesající trend F1 [7, 9, 12, 16]. Druhý formant měl jak tendence klesající [9, 12, 16], tak rostoucí [7]. Metodiky zkoumání jednotlivých studií jsou popsány v kapitole 3.

Formanty jsou používány v širokém rozsahu řečových aplikací (př. parametry pro rozpoznávací systémy, identifikaci mluvčího, při studiu fonetiky).

VSA (vowel space area)

VSA příznak je používán jako akustická míra dysartrie v řeči [24]. Jde o výpočet plochy, která je dána formanty F1, F2 samohlásek /a/, /e/, /i/, /o/, /u/. Výpočet trojúhelníkové plochy pro vokály /a/, /i/, /u/ probíhá podle vzorce 6.4 [23].

$$VSA = \left| \frac{(F1_i \cdot (F2_a - F2_u) + F1_a \cdot (F2_u - F2_i) + F1_u \cdot (F2_i - F2_a))}{2} \right| [\%] \quad (6.4)$$

VAI (vowel articulation index)

Artikulační index samohlásek se získává ze tří samohlásek /a/, /i/, /u/. Výpočet je popsán vzorcem 6.5. Tento příznak je používán pro diagnózu Parkinsonovy choroby.

$$VAI = \frac{F1_a + F2_i}{F1_i + F1_u + F2_u + F2_a} [-] \quad (6.5)$$

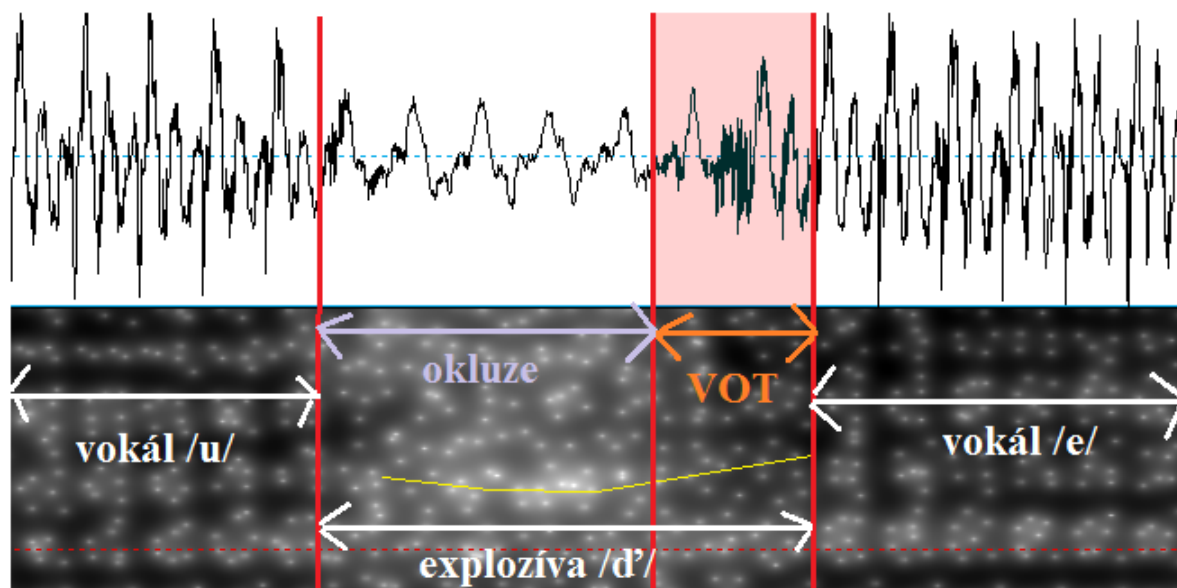
F2i/F2u

Příznak F2i/F2u je dán poměrově. Dokáže dobře rozlišit zdravého člověka od nemocného. Je počítán pouze ze dvou samohlásek a jednoho formantu, proto je VSA efektivnější k odhalování abnormalit [24].

Počátek znělosti VOT (voice onset time)

Parametr počátek znělosti bývá zkoumán na sekvenci vokál – explozíva – vokál nebo jen na zkrácené sekvenci explozíva – vokál. Explozívy jsou podrobně popsány v kapitole 4. Jejich artikulace probíhá ve dvou fázích: okluze neboli závěr a exploze neboli uvolnění závěru. V akustických analýzách jsou měřeny obě doby, doba trvání okluze i doba trvání exploze a jejich poměry k celkové době explozívy. Doba trvání exploze lze definovat jako dobu nástupu hlasivkového tónu a v zahraniční literatuře je označována jako VOT (počátek znělosti). Jedná se o časový úsek mezi uvolněním závěru explozívy a počátkem vibrace hlasivek následující samohlásky. Tento parametr bývá uváděn v souvislosti s rychlostí řeči [14, 15, 23].

Na obr. 6.3 je vykreslena sekvence VEV a jsou v něm označeny jednotlivé segmenty. Díky souvislosti s rychlostí řeči je důležité v analýze měřit doby okluze, exploze ale i doby trvání sousedních vokálů. Hranice mezi jednotlivými segmenty jsou udělány ručně na základě poslechu. Doba exploze je typická výbuchem. Okluze je spíše tupý zvuk. V této práci je analyzována znělá souhláska /dʰ/, a proto se nedají hranice určovat podle přítomnosti či nepřítomnosti základního hlasivkového tónu. Částečně se dá řídit podle intenzity, kde počátek tvoření závěru je dán poklesem intenzity a naopak při explozi intenzita začíná narůstat.



Obr.6.3: Sekvence VEV /u-dʰ-e/ a rozdělené segmenty (žlutě zobrazená intenzita).

Počátek znělosti VOT v rámci podélné studie zkoumali Decoster a Debruyne [1] pro mužský hlas a Košturiaková [14] pro ženský hlas. V prvním případě šlo o zkoumání VOT ve slabikách /pa/ a /ka/. Bylo zjištěno, že s rostoucím věkem roste parametr VOT u obou exploziv. U slabiky /ka/ byl VOT delší než u /pa/. V druhém případě byl zkoumán VOT ze slova /dítě/. Naměřila doby trvání fází explozivy /t/, dobu exploze i okluze a doby sousedních vokálů /i/ a /e/. Analýza byla provedena na celkem pěti slovech /dítě/, a sice dvakrát bylo řečeno v roce 1983, jednou v roce 1988 a dvakrát v roce 2011. Výsledkem její analýzy byly zjištění, že doba exploze, neboli VOT u explozivy /t/ se s rostoucím věkem zkracuje. Doby trvání okluze a sousedních vokálů se měnily skokově, rostly i klesaly, neměly trend v rámci rostoucího věku.

V příčných studiích zkoumal VOT Torre a Barlow [9] na mužských i ženských hlasech. Výsledky potvrdili závislost parametru VOT také na pohlaví. Počátek znělosti byl měřen na 12 segmentech CVC (konsonanta – vokál – konsonanta) pro explozivy /p/, /t/, /k/, /b/, /d/, /g/. U mužů doba exploze (VOT) s rostoucím věkem klesal. U ženského hlasu klesal VOT pouze pro souhlásky /b/ a /d/. U ostatních souhlásek /p/, /t/, /k/, /g/ se VOT prodlužoval s věkem. Explozivy /b/ a /d/ jsou znělé, proto by se dalo předpokládat, že počátek znělosti bude u ženského hlasu klesat u znělých exploziv.

Schötz uvedla ve své analýze [7], že se s rostoucím věkem prodlužují segmenty slov. Neuvedla ale výsledky změn v rychlosti řeči. V díle [8] provedla rešerši předchozích studií věkově závislých parametrů a uvedla, že rychlost řeči roste s věkem v souhláskách a zejména v explozivách. Z toho vyplývá, že VOT se zkracuje.

Machač [15] se zabýval dobou trvání českých exploziv jako takových. Porovnával délky exploziv mezi sebou. Předcházející studie zkoumaly doby trvání explozí a okluzí na čtených projevech řeči, které bývají spojeny s pomalejším tempem promluvy a tím pádem prodlužováním segmentů. Autor zkoumal doby trvání na sekvenci VEV ze spontánních i čtených projevů řeči. Dosáhl kratší délky explozivy. Doba trvání souhlásky /d/ v jeho experimentu byla 74,6 ms (okluze + exploze). V práci uvedl i doby trvání konsonanty /d/ od jiných autorů, a sice 120, 154 a 130 ms.

Pro experimentální ověření věkové závislosti počátku znělosti u ženského hlasu, pro explozivu /d/ se předpokládá, že počátek znělosti i doba trvání okluze se bude s věkem zkracovat. Na sekvenci VEV se bude ověřovat i délka sousedních vokálů.

6.3 Prozodické vlastnosti řeči

Fonační a artikulační analýza se dívají na řeč jako na posloupnost jednotlivých zvuků – akustických segmentů. Jedná se o popis segmentální. Předmětem zájmu je význam řeči. Mluvená řeč obsahuje kromě významu také prozodickou informaci. Jde o vlastnosti řečového signálu, které souvisí hlavně s frekvencí základního hlasivkového tónu, intenzitou a časováním. Mezi aspekty řeči spojené s časováním patří rytmus (přízvuk) a rychlost promluvy. Prozodické vlastnosti se vztahují k větším řečovým jednotkám (úroveň slabik, slov, celých vět) spíše než k samotným segmentům. Jde o popis suprasegmentální. Prozodie se při komunikaci používá k doplnění fonémické informace o jiné lingvistické aspekty, nese

význam o druhu mluvené věty (tázací, oznamovací), o emočním stavu řečníka (rozčilení, radost, smutek) [18].

Suprasegmentální vlastnosti jsou důležité z hlediska syntézy řeči. Ovlivňují přirozenost. Klíčové jsou pro rozpoznávání řečníka, protože jedinečnost projevu člověka spočívá v tom, jak rychle mluví nebo jakým způsobem mění výšku hlasu [18].

6.3.1 Rychlost řeči

Rychlost řeči (tempo) se váže na celkovou rychlost vyslovení promluvy. Většinou se vyjadřuje počtem slabik za sekundu. Rychlost řeči ovlivňuje trvání (délku) jednotlivých segmentů řeči. Slabika se skládá z dalších segmentů, které ovlivňují délku celé slabiky. Změny trvání slabik jsou ovlivněny trváním samohláskových segmentů ve slabice mnohem více než trváním souhlásek [18].

Tempo mluvení je ovlivňováno různými faktory: projevem řeči (čtený, spontánní), emocionálním stavem řečníka (rozčilení, smutek), délkou pauz v promluvě, způsobem artikulace (pečlivá, ledabylá), přízvukem. Rychlost řeči je vázána i na konkrétního mluvčího.

Řada studií se rozchází ve výsledcích změn rychlosti řeči s věkem. Často není tento parametr ani hodnocen, pouze jsou hodnoceny změny délky segmentů. Tempo závisí také na parametru VOT a pauzách v promluvě.

Z rešerše se v rámci podélné studie zabývala rychlostí řeči pouze Košturiaková [14]. Analyzovala délku trvání jednotlivých fonémů i samotného slova /sama/ a /dítě/. Zjistila, že celková délka slova se s věkem zkracuje (zkracuje se doba exploze), ale neprokázala, že by se rychlost řeči zrychlovala nebo zpomalovala.

Schötz [7] ve své studii zjistila, že se délka slov prodlužuje s věkem. Ve studii [8] napsala, že rychlost řeči se zrychluje s rostoucím věkem v souhláskách a explozích. To znamená, že se zkracuje počátek znělosti (VOT) a prodlužují se samohlásky nebo pauzy.

Báňa [13] zkoumal rychlost řeči (počet slabik za sekundu) v šesti větech u různých věkových skupin. v experimentu zjistil, že rychlost řeči se zpomaluje s věkem.

Janda [16] analyzoval rychlost řeči u dětí do věku 15 let na dětské říkánce. Spolu s rychlostí analyzoval i délky pauz. Zjistil, že rychlost řeči se zvětšuje a doby pauz se zmenšují. V této věkové skupině se předpokládalo, že děti ve věku 13 až 15 budou mít lepší vyjadřovací schopnosti a proto bude méně pauz v promluvě a celkové tempo bude rychlejší. Ve svém experimentu potvrdil nárůst rychlosti v závislosti na zkracování pauz v promluvě.

6.3.2 Podíl pauz v promluvě

Dalším faktorem, který ovlivňuje nejen časování, ale i srozumitelnost řeči, jsou pauzy. Délka pauz může být různá a s rostoucím věkem se může měnit. Procentuální zastoupení pauz v promluvě lze vypočítat podle následujícího vzorce 6.6 [16].

$$pp = \frac{l_t - l_a}{l_t} \cdot 100 [\%], \quad (6.6)$$

kde l_t je celková doba promluvy, l_a je suma délek jejích artikulovaných úseků.

7. Analýza věkové závislosti

Na základě rešerše byla v této práci ověřována věková závislost vybraných akustických parametrů na databázi ženského hlasu. Byla vytvořena databáze filmových záznamů herečky Ivany Chýlkové. Ze záznamů byly nastříhány její promluvy, od slov až po velká souvětí. Ty byly dále zpracovávány pomocí softwaru Praat. Byly analyzovány vybrané hlasové charakteristiky. Z hlediska fonačního byla analyzována základní frekvence (F0), parametr frekvenční variability základního hlasového tónu jitter, amplitudové variability shimmer. Z pohledu artikulačních vlastností byly zkoumány první dvě formantové frekvence (F1, F2), vokální trojúhelník, velikost vokálního prostoru (VSA), artikulační index (VAI), kombinace poměrů formantů vokálů /a/, /i/, /u/, počátek znělosti (VOT). Z prozodických vlastností řečového signálu byla zkoumána rychlost řeči a procentuální zastoupení pauz v promluvě.

7.1 Databáze

Pro účely analýzy věkově závislých akusticko-fonetických parametrů hlasu byla vytvořena databáze nahrávek ženského hlasu, herečky Ivany Chýlkové. Filmové záznamy z období let 1985-2014 byly získány z internetových stránek www.hellspy.cz a www.youtube.cz ve formátu avi. Seznam filmů, seriálů i rozhovorů je uveden v příloze C.

Pro zpracování zvukových souborů bylo potřeba získat formát wav, proto byly nahrávky převedeny z formátu avi na wav pomocí nástroje Any Video Converter.

Z filmů byly extrahovány různě dlouhé promluvy herečky pomocí software Cool Edit Pro 2.0. Celkem bylo extrahováno 962 záznamů (vět). Seznam vět pro jednotlivé roky je uveden v příloze D. V tab. 7.1 je uveden počet promluv pro jednotlivé roky analyzovaného období 1985 – 2014. Dále byly zvukové signály zpracovány v Praatu.

Pro analýzu byly vybrány co nejklidnější promluvy, bez velkého podílu šumu způsobeného hudbou v pozadí nebo emočním vypjetím herečky díky dané roli.

Rok	Věk	Počet záznamů	Rok	Věk	Počet záznamů	Rok	Věk	Počet záznamů
1985	22	39	1994	31	17	2007	44	105
1986	23	18	1997	34	18	2008	45	45
1987	24	14	1998	35	76	2009	46	5
1988	25	3	1999	36	27	2011	48	125
1989	26	94	2000	37	45	2012	49	2
1990	27	35	2001	38	27	2013	50	3
1992	29	70	2002	39	64	2014	51	10
1993	30	57	2005	42	63			
Celkem záznamů (vět)								962

Tab.7.1: Databáze extrahovaných záznamů z filmů herečky.

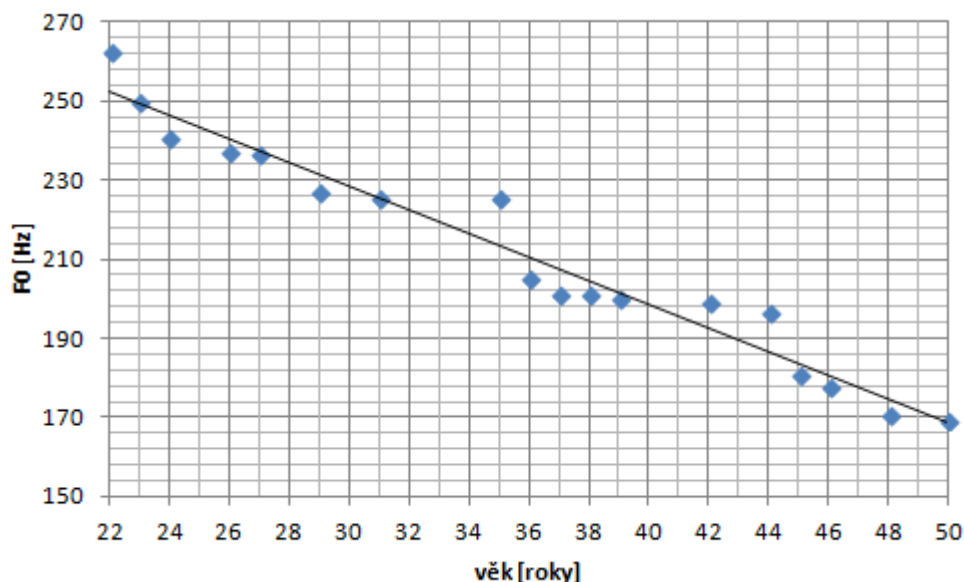
7.2 Základní frekvence

Základní frekvence byla analyzována z prodloužených fonací samohlásky /a/ z jednotlivých záznamů v období let 1985 – 2013 hereččiny filmové kariéry. V každém roce byl použit různý počet prodloužených fonací vokálu /a/. Zkoumány byly všechny samohlásky bez rozdílu pozice samohlásky ve slově. Různý počet byl dán počtem celkových nahrávek v daném filmovém období, a zda šlo o hlavní či vedlejší roli herečky. V tabulce 7.2 je uveden počet vokálů použitý pro získání průměrné hodnoty základního hlasivkového tónu v jednotlivých letech jako věkově závislého parametru hlasu. Každá samohláska je dána výčtem hodnot frekvencí rozdílné délky trvání. Analyzovány byly pouze hodnoty bez velkého rušení hudebním pozadím a přehnaných emocí. Hodnoty F0 byly vypočteny detekčním algoritmem v Praatu označením vybraného segmentu.

Rok	Věk	Počet /a/	Rok	Věk	Počet /a/	Rok	Věk	Počet /a/
1985	22	14	1994	31	5	2005	42	5
1986	23	8	1998	35	6	2007	44	5
1987	24	5	1999	36	7	2008	45	3
1989	26	8	2000	37	5	2009	46	2
1990	27	2	2001	38	5	2011	48	7
1992	29	5	2002	39	7	2013	50	1

Tab. 7.2: Počet analyzovaných prodloužených fonací vokálu /a/.

V grafu 1 je znázorněn průběh vypočtených průměrných hodnot F0 z hodnot frekvencí základního hlasivkového tónu z daných záznamů v tab.7.2.



Graf 7.1: Průběh F0 ženského hlasu z prodloužené fonace /a/ s rostoucím věkem.

Z grafu 7.1 je vidět, že F0 s rostoucím věkem klesá u ženského hlasu.

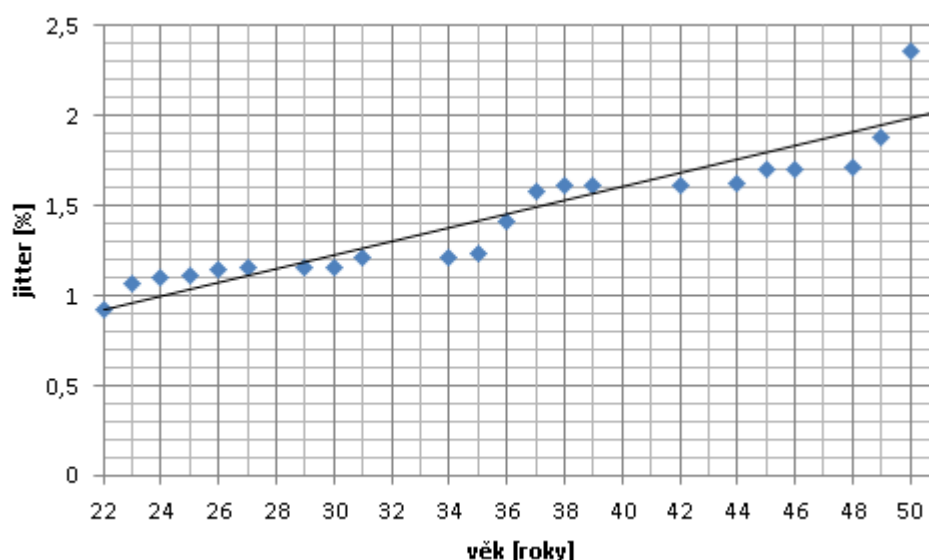
7.3 Jitter a shimmer

Parametr frekvenční variability základní frekvence a amplitudové variability byly určeny z prodloužených fonací vokálu /a/ a v letech, kde byl nedostatek dat pouze ze základní fonace /a/ v Praatu. Jitter i shimmer určují kvalitu fonace, ale jsou také hodně citlivé na kvalitu použitých nahrávek pro analýzu. Díky tomu byly některé nahrávky vyřazeny a průběh průměrných hodnot obou parametrů byl počítán z menšího počtu dat. V tab. 7.3 je uveden počet hodnot lokálního jitteru/shimmeru pro dané roky. Jde pouze o jednu procentuální hodnotu. Daná hodnota už v sobě neobsahuje výčet dalších hodnot jako v případě základní frekvence, viz tab. 7.2

Rok	Věk	Počet hodnot	Rok	Věk	Počet hodnot	Rok	Věk	Počet hodnot
1985	22	5	1994	31	2	2007	44	12
1986	23	2	1997	34	2	2008	45	2
1987	24	2	1998	35	6	2009	46	2
1988	25	1	1999	36	4	2011	48	33
1989	26	11	2000	37	5	2012	49	1
1990	27	6	2001	38	2	2013	50	1
1992	29	11	2002	39	4	2014	51	6
1993	30	11	2005	42	8			

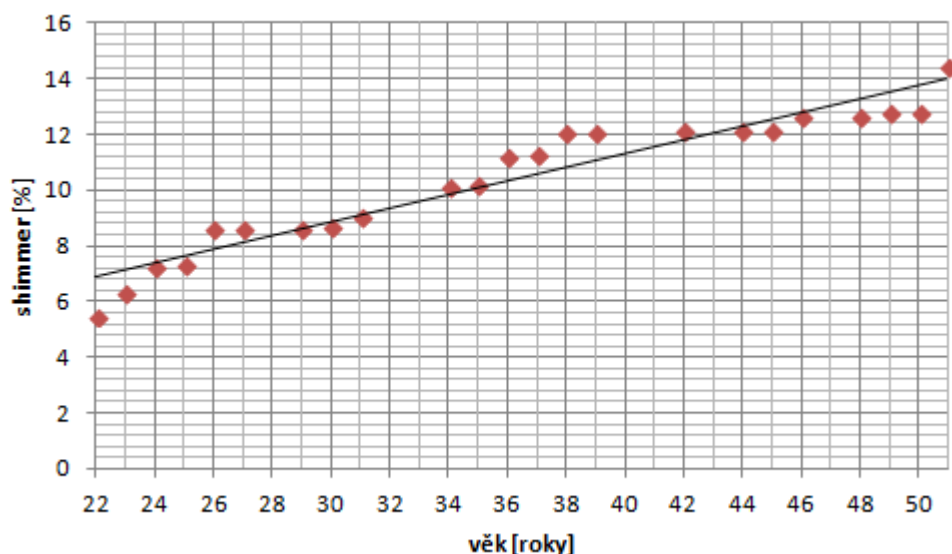
Tab. 7.3: Počet hodnot pro výpočet jitteru.

V grafu 7.2 je znázorněn průběh průměrné hodnoty parametru jitter s rostoucím věkem proložený lineární regresní přímkou. Přímka slouží spíše pro znázornění trendu, zda jde o rostoucí či klesající.



Graf 7.2: Průběh parametru jitter s rostoucím věkem.

V grafu 7.3 je znázorněn průběh průměrné hodnoty parametru shimmer s rostoucím věkem opět proložený lineární regresní přímkou.



Graf 7.3: Průběh parametru shimmer s rostoucím věkem.

7.3.1 Hodnocení variabilit F0

Z grafů 7.2 a 7.3 vyplývá, že frekvenční i amplitudová variabilita s rostoucím věkem roste. Předpokladem pro analýzu těchto parametrů bylo, že s rostoucím věkem narůstá kolísání jak frekvence, tak amplitudy. Řada studií zkoumající věkovou závislost potvrzuje růst s rostoucím věkem a vysvětluje toto chování oslabením hlasivkové pružnosti či celkovou únavou hlasu.

Hodnota jitteru stoupla z 0,92 % ve 22ti letech na 2,36% v 51 letech. U frekvenční variability nedochází k velkým nárůstům s rostoucím věkem. Dále se pro analýzu parametru používají vzorkovací frekvence vyšší než 22 kHz, ale v této analýze byly všechny řečové signály převzorkovány na stejnou vzorkovací frekvenci, a sice na 16 kHz, která je dostačující pro zpracování řečového signálu [17].

Hodnota shimmeru je vyšší než hodnota jitteru. V analýze se pohybuje zhruba z 5ti % na 15% během 30ti let. Jde o změnu amplitudy, kde dochází k většímu rozdílu než ve změnách frekvence.

V datech mohlo dojít ke zkreslení díky kvalitě použitých nahrávek, proto pro věrnější výsledky by bylo třeba získat větší množství hodnot.

7.4 Formantové frekvence

V této diplomové práci byly analyzovány první dva formanty z prodloužených fonací samohlásek /a/, /e/, /i/, /o/, /u/. Tyto segmenty byly extrahovány ze všech záznamů. Počet záznamů uveden v tab.7.4. Dále byl pro každou samohlásku upraven frekvenční rozsah v Praatu v záložce *Formant Advance settings* pro detekci formantových frekvencí, jejichž kombinace jsou typické pro každou samohlásku.

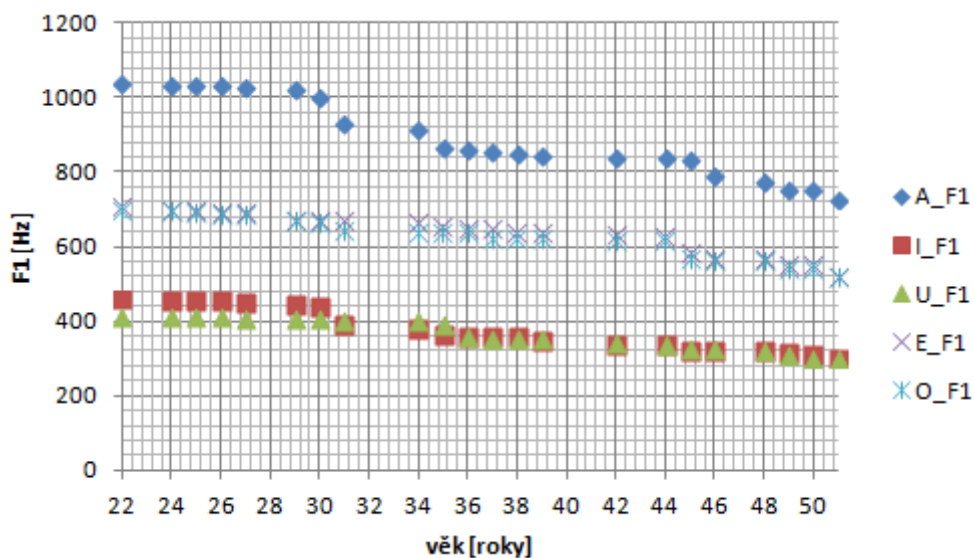
Pro detekci /u/ byly hledány dva formanty do 1000 Hz. Pro /i/ byly nastaveny dva formanty do 3000 Hz. Pro /a/ byly nastaveny tři formanty do 3000 Hz. Detekce formantových pásů byla daleko úspěšnější než při defaultním nastavení hledání pěti formantových pásů do 5500 Hz.

V každém roce záznamu byl pro výpočet průměrné hodnoty použit rozdílný počet prodloužených fonací vokálů. V tabulce 7.4 je uveden počet jednotlivých samohlásek pro daný rok, a sice v pořadí /a/, /e/, /i/, /o/, /u/. Každá hodnota v sobě obsahuje průběh frekvencí, průměrují se tedy celé rozsahy trvání vokálů.

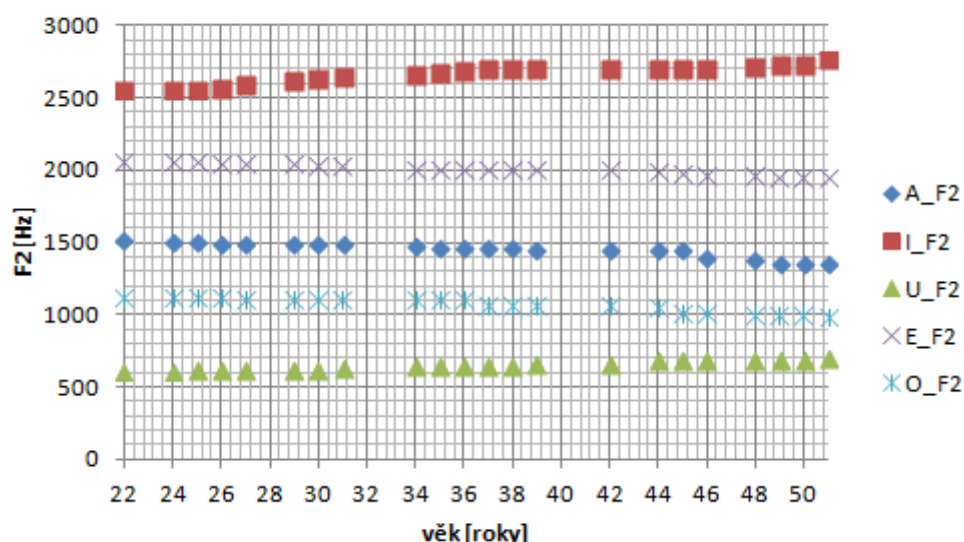
Rok	Věk	Počet /a/e/i/o/u/	Rok	Věk	Počet /a/e/i/o/u/	Rok	Věk	Počet /a/e/i/o/u/
1985	22	5/3/2/3/3	1994	31	3/2/4/1/2	2005	42	6/3/4/1/4
1987	24	2/1/1/1/1	1997	34	6/4/3/2/1	2007	44	6/2/3/2/5
1988	25	2/1/2/2/1	1998	35	8/5/8/3/5	2008	45	3/1/2/1/1
1989	26	5/1/3/3/2	1999	36	8/3/5/3/2	2009	46	1/1/1/1/1
1990	27	2/2/2/1/2	2000	37	6/2/5/2/3	2011	48	8/2/3/4/4
1992	29	6/2/2/1/3	2001	38	2/3/2/1/4	2013	50	2/1/2/1/1
1993	30	2/1/3/1/1	2002	39	6/1/3/1/5	2014	51	4/4/4/1/1

Tab. 7.4: Počet analyzovaných prodloužených fonací vokálů /a/, /e/, /i/, /o/, /u/.

Průběhy F1, F2 s rostoucím věkem pro samohlásky vykresleny v grafech 7.4, 7.5.



Graf 7.4: Průměrné hodnoty F1 s rostoucím věkem u ženského hlasu ve věku 22 – 51 let.



Graf 7.5: Průměrné hodnoty F2 s rostoucím věkem u ženského hlasu ve věku 22 – 51 let.

7.4.1 Hodnocení formantů F1, F2

Z grafu 7.4 vyplývá, že první formant u všech samohlásek s rostoucím věkem klesá. Tento výsledek potvrzuje většina studií zabývajících se věkovou závislostí prvního formantu.

Pro prodlouženou fonaci vokálu /a/ naměřená hodnota F1 klesá z 1039 Hz na 730 Hz. Tabulkové rozpětí F1 pro /a/ je 700 – 1100 Hz [17]. Pro vokál /e/ naměřená hodnota F1 klesá ze 711 Hz na 523 Hz. Tabulkové rozpětí F1 pro /e/ je 480 – 700 Hz [17]. Pro vokál /i/ naměřená hodnota F1 klesá ze 462 Hz na 304 Hz. Tabulkové rozpětí F1 pro /i/ je 300 – 500 Hz [17]. Pro vokál /o/ naměřená hodnota F1 klesá ze 700 Hz na 522 Hz. Tabulkové rozpětí F1 pro /o/ je 500 – 700 Hz [17]. Pro /u/ naměřená hodnota F1 klesá ze 412 Hz na 303 Hz. Tabulkové rozpětí F1 pro /u/ je 300 – 500 Hz [17].

Z grafu 7.5 je vidět, že druhý formant s věkem klesá pouze u samohlásek /a/, /e/, /o/. U samohlásek /i/, /u/ dochází s rostoucím věkem k nárůstu F2.

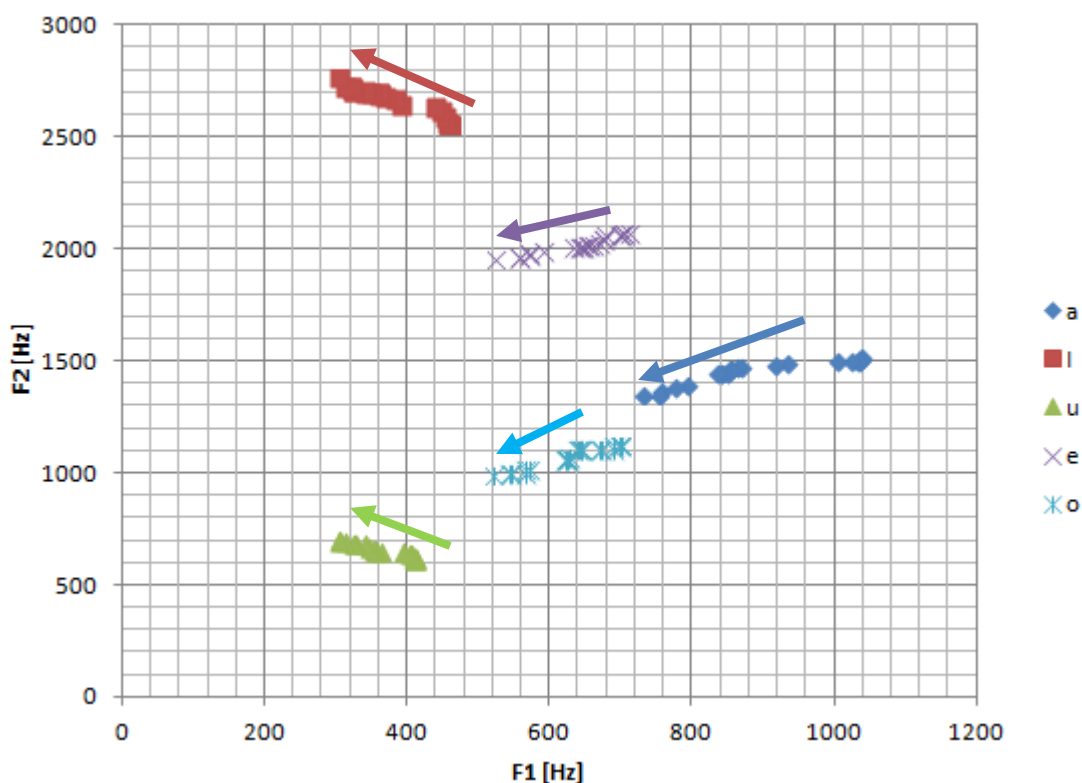
Naměřená hodnota F2 pro /a/ klesá z 1513 Hz na 1350 Hz. Tabulkové hodnoty F2 pro /a/ jsou 1100 – 1500 Hz. Naměřená hodnota F2 pro /e/ klesá z 2066 Hz na 1957 Hz a tabulkově se vokál /e/ nachází v rozmezí 1500 – 2100 Hz. Naměřená hodnota F2 pro vokál /o/ klesala z 1118 Hz na 989 Hz. Tabulkové hodnoty pro druhý formant /o/ jsou 850 – 1200 Hz.

Naměřená hodnota F2 pro samohlásku /i/ vzrostla v průběhu třiceti let z 2559 Hz na 2769 Hz. V tabulce se najdou pro druhý formant vokálu /i/ hodnoty 2000 – 2800 Hz. Získaná hodnota F2 pro vokál /u/ vzrostla během třiceti let z 614 Hz na 701 Hz. Tabulkově se F2 této samohlásky pohybuje v rozmezí 600 – 1000 Hz.

Hodnoty obou formantů u samohlásek byly detekovány, po nastavení mezí pro hledání formantových pásů v Praatu, bez problémů. Kvalita nahrávek zásadně neovlivnila získané hodnoty pro všechny analyzované samohlásky.

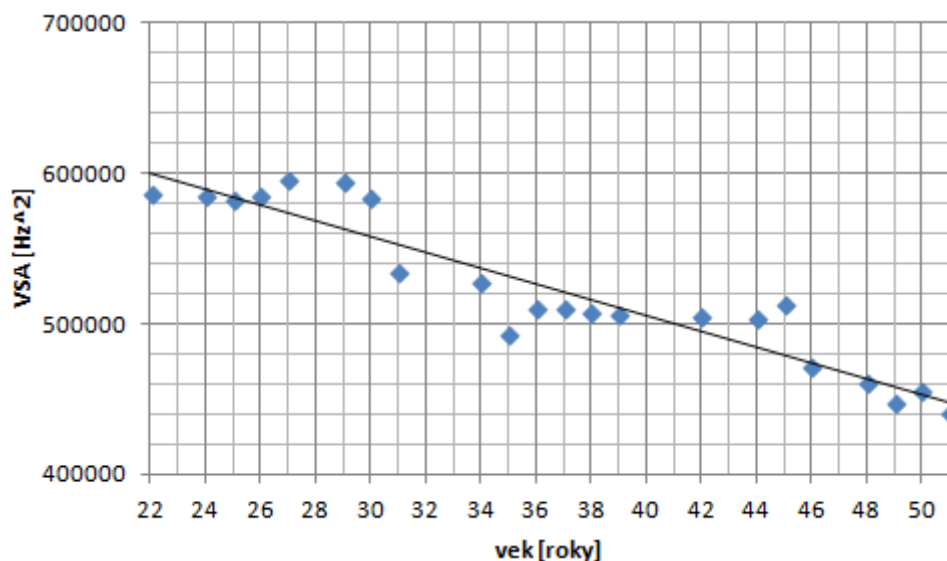
7.5 Vokální trojúhelník a parametr VSA

Samohlásky tvoří vokální trojúhelník v závislosti na kombinaci F1 a F2. Vrcholy trojúhelníku tvoří vokály /i/, /a/, /u/. V předchozí kapitole byly znázorněny průběhy F1 i F2 v čase. V této kapitole jsou znázorněny změny pozic samohlásek ve vokálním trojúhelníku, viz graf 7.6. Šipkami je znázorněn posun středních hodnot F1 a F2 daných samohlásek s rostoucím věkem v průběhu věku 22 – 51 let.



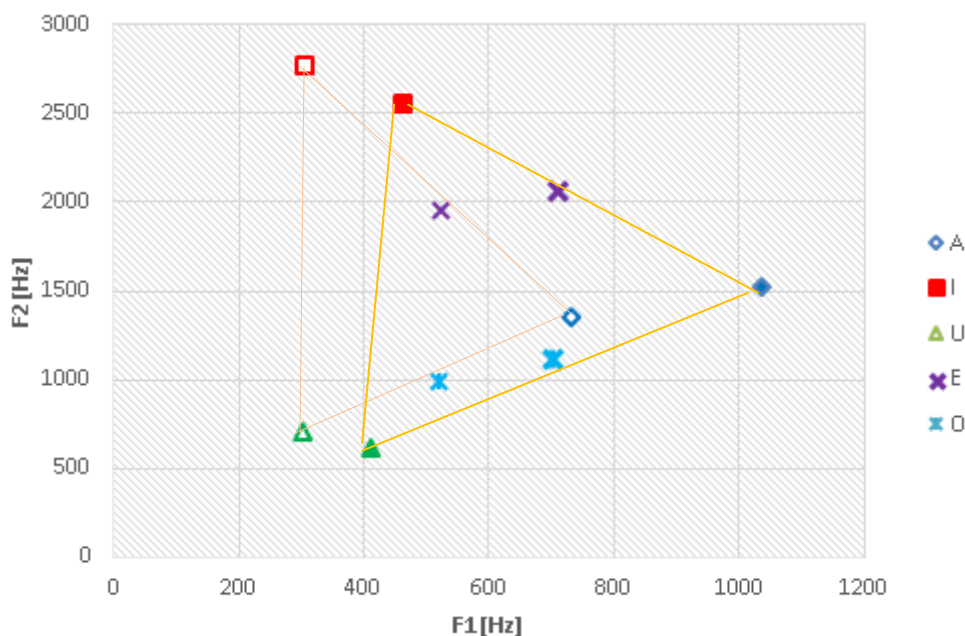
Graf 7.6: Změny pozic středních hodnot F1, F2 daných samohlásek v průběhu 22 – 51 let.

V grafu 7.7 je znázorněn průběh vypočítaného parametru VSA podle vzorce 6.4.



Graf 7.7: Průběh parametru VSA s rostoucím věkem.

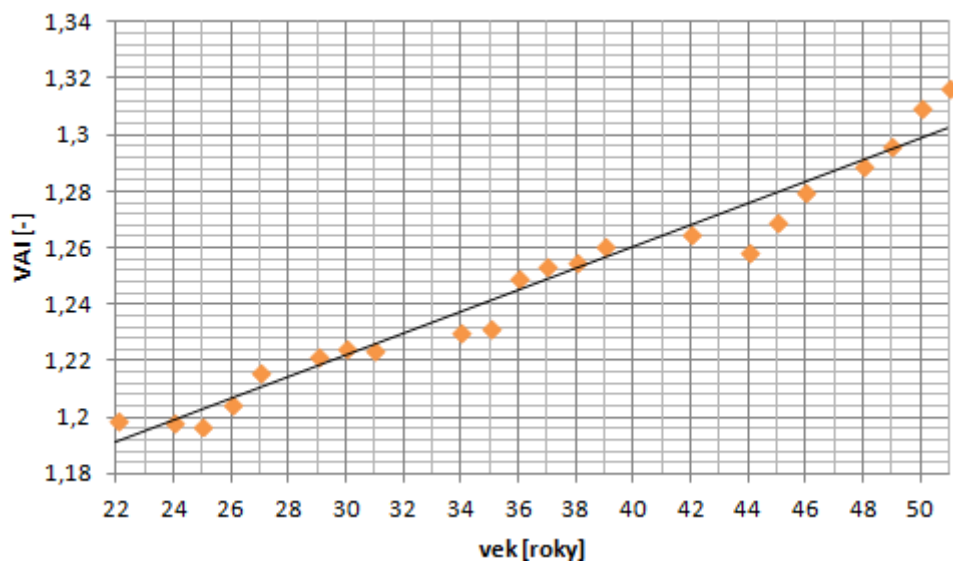
Z grafu 7.6 vyplývá, že F1 u všech vokálů klesal (posun doleva). F2 klesal u /a/, /o/, /e/ a rostl u /i/ a /u/. Na základě toho bylo předpokládáno, že se bude zmenšovat plocha vokálního prostoru s rostoucím věkem. V grafu 7.7 byl potvrzen klesající průběh parametru (plocha vokálního prostoru se s věkem zmenšuje). Pro názornost jsou c grafu 7.8 vykreslené vokalické trojúhelníky pro věk 22 (tmavší oranžový) a 51 let (světlý růžový). Při prvním pohledu je vidět, že plocha předního trojúhelníku (pro věk 22) je větší než druhá plocha.



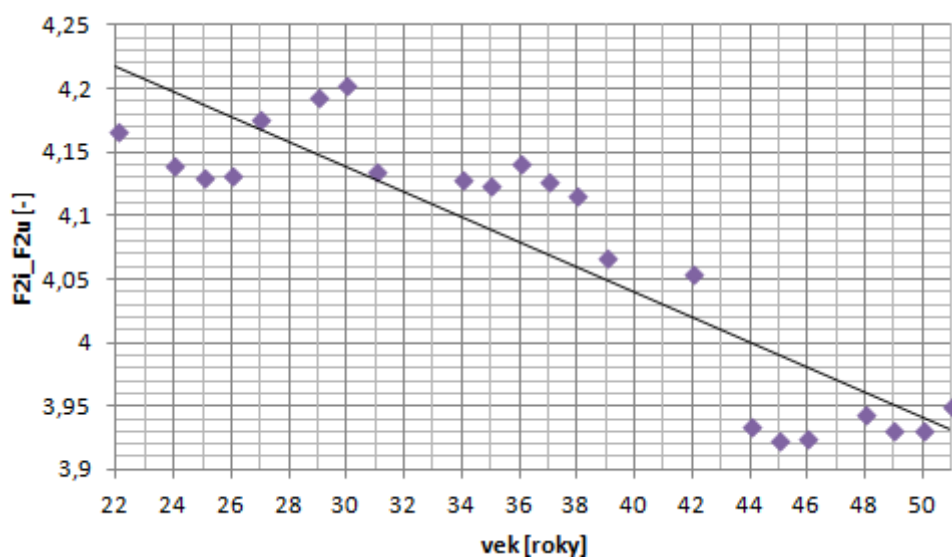
Obr.7.8: Vokální trojúhelníky ve 22 (tmavě oranžový) a 51 letech (světlý).

7.6 VAI a poměr $F2i/F2u$

VAI byl vypočítán v Matlabu podle vzorce 6.5 a jeho průběh s věkem je znázorněn v grafu 7.9. Poměr $F2i/F2u$ byl vypočten poměrem druhých formantů daných vokálů a jeho průběh je znázorněn v grafu 7.10.



Graf 7.9: Průběh parametru VAI s rostoucím věkem.



Graf 7.10: Průběh poměrů $F2i/F2u$ s rostoucím věkem.

7.6.1 Hodnocení VAI a $F2i/F2u$

Z grafu 7.9 vyplývá, že s rostoucím věkem hodnota artikulačního indexu narůstá. Vypočtené hodnoty VAI mohou být použité v databázi zdravých jedinců pro rozpoznávání Parkinsonovy choroby.

Trend poměru druhých formantů samohlásek /i/ a /u/ s rostoucím věkem klesá. Druhé formanty u /i/ a /u/ s rostoucím věkem rostou. Z grafu 7.10 jsou vidět změny, kdy roste $F2$ u vokálu /i/ rychleji a kdy je větší $F2$ u /u/, střídání klesajících nebo rostoucích hodnot.

7.7 Počátek znělosti VOT

Počátek znělosti (VOT, doba exploze) byl experimentálně zkoumán s dobou trvání okluze explozivní na databázi slov /udělat/, /nedělat/ z databáze všech záznamů (tab. 7.1, seznam vět v příloze D). Slovo /udělat/ bylo řečeno ve většině zkoumaných letech pořízených záznamů. Doba trvání exploze (počátek znělosti), okluze, celková doba explozivní (okluze + exploze) a doby trvání sousedních vokálů byly změřeny na sekvenci /u-d'-e/, tab. 7.5. Analýza byla doplněna o měření sekvence /e-d'-e/, viz tab. 7.6. Pro analýzu byla použita slova z co nejklidnějších promluv bez emocí a velkého rušení hudbou, okolím. Ale některá slova mohla být řečena důrazněji (rychleji) v rámci dané filmové role. Do zpracování nebyla zahrnuta slova z roku 2014 (věk 51), protože šlo o čtený projev herečky. Ostatní záznamy byly spontánní promluvy z filmů. Machač [15] uvedl, že při čtených promluvách dochází k celkovému prodlužování jednotlivých segmentů explozivní i sousedních vokálů.

Zpracování zvukových nahrávek slov bylo provedeno v Praatu. Určování hranic jednotlivých hlásek bylo provedeno ručně (popsáno v kapitole 6.2.4).

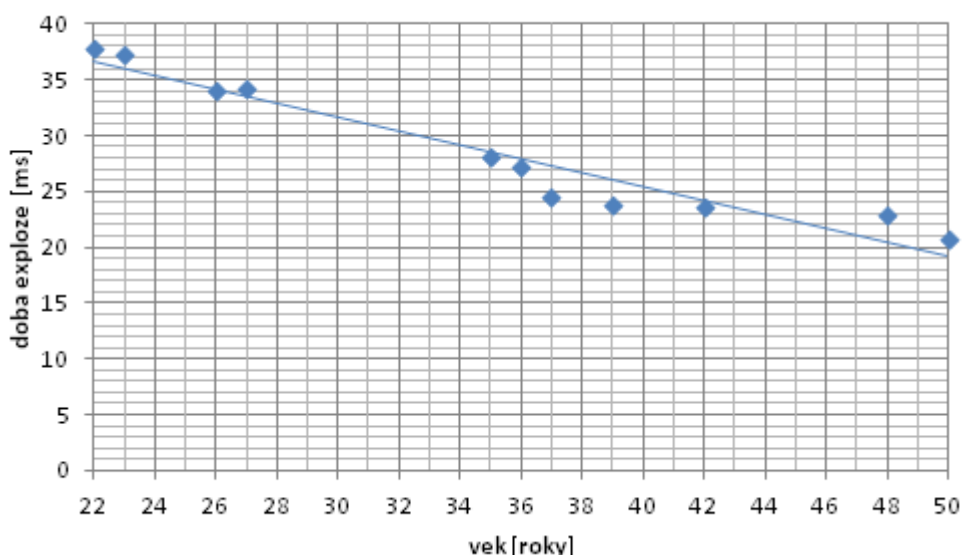
Ruční segmentace je uvedena v příloze D v souboru Praat.Collection. Segmentované obrázky jsou vykreslené na obr. A2 – A20 v obrazové příloze A pro jednotlivé roky.

Počet /u-d'-e/	Věk	Vokál /u/ [ms]	Explozíva /d'/			Vokál /e/ [ms]
			Doba okluze [ms]	Doba exploze [ms]	Celková doba /d'/ [ms]	
1	22	31,6	41,3	37,9	79,2	45,2
1	23	47,6	39,3	37,2	76,5	45,2
3	26	47,1	37,1	30,9	68,0	55,8
	26	49,5	38,4	34,5	72,9	54,1
	26	30,2	38,8	36,8	75,6	45,2
1	27	66,8	37,0	34,3	71,3	72,9
2	29	Vyřazeno z analýzy (rušení)				
	29					
3	35	40,0	26,9	28,5	55,4	38,0
	35	62,5	27,7	26,5	54,2	58,2
	35	31,5	23,5	29,1	52,6	35,5
1	36	46,4	25,8	27,2	52,9	48,8
1	37	44,0	24,2	24,4	48,6	38,2
1	39	32,4	23,4	23,8	47,2	44,5
2	42	54,6	23,1	23,6	46,7	57,5
	42	Vyřazeno z analýzy (emoce)				
5	48	44,0	22,0	23,1	45,1	34,8
	48	45,4	22,5	23,6	46,1	70,0
	48	43,8	22,7	23,6	46,3	35,6
	48	47,2	22,5	23,5	46,0	42,8
	48	30,3	21,1	20,6	41,7	97,9
1	50	54,6	20,1	20,8	40,9	55,4
			Klesá	Klesá	Klesá	
22	Celkem slov /udě/ (-lat, -lám, -láš, -lal, -lala, -lali, -laly)					
19	Počet slov po vyřazení nepoužitelných slov (rušená)					

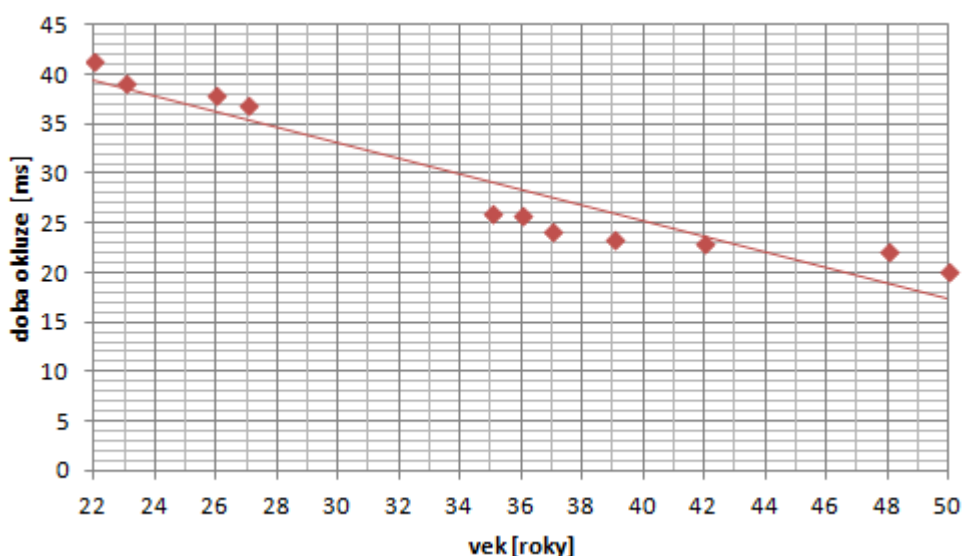
Tab. 7.5: Naměřené doby trvání vokálů, explozívy v sekvenci /u-d'-e/.

Doba exploze (počátek znělosti) i doba okluze se s věkem zkracují. Trend exploze je v grafu 7.11, trend okluze v grafu. 7.12. Oba trendy jsou klesající. Celkově se doba explozivní s rostoucím věkem zkracuje. Rychlost řeči v explozivách roste.

Doby trvání sousedních samohlásek bylo ovlivněno rychlostí vyslovení. Slova jsou v průběhu let vyslovena různou rychlostí. Vyšší rychlost způsobuje zkrácení samohlásek, pomalejší rychlost prodlužování segmentů (samohlásek). Délka vokálů se měnila v závislosti na rychlosti vyslovení. Vliv věku se na prodlužování / zkracování neprojevil.



Graf 7.11 Průběh doby exploze (VOT)



Graf 7.12 Průběh doby okluze (VOT)

Machač [15] naměřil průměrnou dobu explozivní /d'/ 74, 6 ms, z toho závěr trval 49,1 ms a VOT 25,4 ms. Neuvedl ale věk subjektů, u kterých doby trvání měřil. Závěr ku explozi měl v poměru 2:1. V tomto experimentu je doba okluze s dobou exploze přibližně v poměru 1:1. Zde byly zpracovávány ale pouze mluvené projevy na rozdíl od jeho práce, kde byly zkoumány mluvené i čtené, a proto jsou doby trvání kratší.

V tab. 7.6 jsou obdobně doby trvání jednotlivých hlásek v sekvenci /e-d'-e/. Ruční segmentace k dispozici v příloze D v souboru praat.Collection2.

Počet /u-d'-e/	Věk	Vokál /e/ [ms]	Exploziv /d'/			Vokál /e/ [ms]
			Doba okluze [ms]	Doba exploze [ms]	Celková doba /d'/ [ms]	
1	24	52,9	40,0	33,1	73,1	48,9
1	26	39,8	37,7	28,4	66,1	36,8
1	37	50,1	33,3	27,3	60,6	66,0
1	39	50,3	33,2	27,0	60,2	47,0
1	42	69,0	33,0	25,7	58,7	71,0
			Klesá	Klesá	Klesá	
5	Celkem slov /edě/					

Tab. 7.6: Naměřené doby trvání vokálů, explozivy v sekvenci /e-d'-e/.

Doba okluze i exploze (VOT) s rostoucím věkem klesaly a byly v poměru 1:1 jako u sekvence /u-d'-e/. Délky sousedních vokálů už ale trochu naznačují, že by se mohly s věkem prodlužovat. Sekvence /e-d'-e/ se lišila od první sekvence v použitých slovech, která se nacházela většinou na začátku slov (nepředcházela jí žádná předpona až na výjimky /neudělal/). Druhá analyzovaná sekvence měla před prvním vokálem většinou souhlásku /n/ nebo předponu /př/. Částečně pozice mohla ovlivnit délku prvního vokálu.

Prodlužování samohlásek má souvislost s rychlostí vyslovení promluvy. S vyšším věkem spíše herečka některé hlásky „zamumlala“, než že by prodlužovala. Tato domněnka je čistě ze subjektivního hodnocení poslechu ženského hlasu v průběhu třiceti let. Artikulace je horší s věkem, některé hlásky byly dokonce těžko rozlišitelné z nahrávek. Vliv má samozřejmě i kvalita nahrávek.

7.7.1 Hodnocení VOT

Obrázky A1 – A19, průměrné hodnoty dob trvání jednotlivých fází explozivy (tab. B1) ukazují zkracování dob s rostoucím věkem, což potvrzuje výsledky práce [8]. Tím, že se zkracují doby trvání explozivy, roste rychlost řeči v tomto segmentu. Na experimentálních datech nebyla potvrzena souvislost věku s prodlužováním sousedních vokálů explozivy. Proměnlivá doba trvání vokálů je spojena s celkovou rychlostí promluvy a ostatními prozodickými vlastnostmi řeči (důraz, intenzita) a hereckou rolí.

7.8 Rychlost řeči

Z prozodických vlastností řečového signálu byla v této práci zkoumána věková závislost rychlosti řeči. Jednotkou tempa je počet slabik za sekundu. Každá slabika se skládá z dalších segmentů, které ovlivňují délku celé slabiky. Změny trvání slabik jsou ovlivněny trváním samohláskových segmentů ve slabice mnohem více než trváním souhlásek, proto bylo důležité vybrat věty s podobným (stejným) složením samohlásek. Experimentální data z filmů jsou omezená danou hereckou rolí, proto nelze najít stejné věty v jednotlivých rocích získaných promluv.

Z analýzy rychlosti řeči byla vyloučena data z roku 2014. Jednalo se o čtený projev herečky. Čtení bývá spojováno s pomalejší rychlostí, a proto data z tohoto roku nebyla použita pro analýzu rychlosti řeči.

Při zkoumání počátku znělosti bylo zjištěno, že se doba trvání explozivů s rostoucím věkem zkracuje. To má za následek zvyšování rychlosti v těchto segmentech. Schötz dokonce tvrdila, že rychlost řeči se zvyšuje s rostoucím věkem u všech souhlásek [8].

V získaných datech nebylo možné měřit rychlost celkové promluvy, protože složení vět je nehomogenní. Bylo vybráno několik slov, u nichž byla porovnávána doba trvání.

Rychlost byla měřená na stejných slovech v různých letech, kdy byla vyslovena s podobnými emocemi. Slova byla vybrána náhodně podle subjektivního hledání v seznamu vět. V tab. 7.7 jsou výsledky pro slovo /dobře/ a první dvě slabiky slova /potřebovat/, u nichž se podařilo najít trend s rostoucím věkem.

Slovo /do bře/	Doby trvání slabik [ms]		/potřebovat/	Doba trvání [ms]
věk	/do/	/bře/	Věk	/potře/
37	7,4	17,6	23	25,3
39	8,0	17,7	26	24,4
48	8,4	18,1	38	21,4
50	11,4	21,7	39	16,0
			44	15,5
	Prodlužuje se	Prodlužuje se		Zkracuje se

Tab. 7.7: Naměřené doby trvání slabik.

7.8.1 Hodnocení rychlosti řeči

Filmové záznamy nejsou vhodným materiálem pro analýzu rychlosti řeči ani pro podíl pauz v promluvě. Není možné nalézt stejné, skoro ani podobné, věty, na nichž by se dal parametr zkoumat. Částečně se dá hodnotit rychlost zkoumáním počátku znělosti, jak již bylo výše popsáno. Tab. 7.7 ukazuje, že některé slabiky se prodlužují s rostoucím věkem a některé se zkracují. Tato data nelze hodnotit z prozodické stránky, jsou příliš zaujatá vlivem filmové role. Jediného, čeho bylo experimentálně ověřeno, že rychlost řeči se zvyšuje v explozivě /d'/.

7.9 Zhodnocení výsledků

Základní frekvence byla analyzována na prodloužené fonaci samohlásky /a/. V práci byla potvrzena věková závislost tohoto parametru, který s rostoucím věkem klesá. Směrodatná odchylka základě analýzy filmových záznamů nebyla hodnocena. Díky kvalitě záznamů i hereckému nasazení byla vidět ze záznamů práce s hlasem, proto se směrodatná odchylka F0 nedala hodnotit. Neměla žádný trend, jednalo se o skokové změny.

Frekvenční variabilita jitter a amplitudová variabilita shimmer jsou parametry kvality fonace. S rostoucím věkem dochází k narůstání těchto nestabilit F0. V práci bylo potvrzeno, že oba parametry s rostoucím věkem rostou. Byly hodnoceny ale na malém množství dat, protože jsou citlivé na kvalitu nahrávek a část dat nemohla být pro analýzu použita.

Věková závislost prvních dvou formantů F1 a F2 byla potvrzena na prodloužených fonacích samohlásek /a/, /e/, /i/, /o/, /u/. F1 u všech vokálů s rostoucím věkem klesá. F2 klesá u samohlásek /a/, /e/, /o/. Naopak u samohlásek /i/, /u/ F2 s rostoucím věkem roste.

Z naměřených formantů byl vytvořen vokalický trojúhelník pro všechny roky. Byla vypočítán parametr VSA. Velikost vokálního prostoru se s rostoucím věkem snižuje. Dále byl počítán z formantů artikulační index, který s rostoucím věkem roste. V této práci byl hodnocen i poměr druhých formantů vokálů /i/ ku /u/. Ten s rostoucím věkem klesá.

Počátek znělosti byl analyzován ze sekvence VEV, a sice segmentu /u-d'-e/ a /e-d'-e/. S rostoucím věkem se zkracovaly doby trvání okluze i exploze (VOT). U sousedních vokálů nebyla potvrzená věková závislost, změny v době trvání jsou spojené s celkovou rychlostí promluvy. Rychlost řeči ve zkoumané explozivě se s věkem zrychluje.

Výzkum rychlosti řeči celých promluv z filmových záznamů se ukázal jako nesprávný. Databáze obsahovala omezené promluvy, kde nebyla možnost najít shodnost či velmi blízká podobnost vět či slov. Každý záznam měl jiné složení slov, slabik, hlásek a nelze sjednotit délky všech slabik. Částečná analýza proběhla na několika vybraných slovech (slabikách). Některé slabiky se s věkem prodlužovaly, některé zkracovaly. Nelze říci, zda jde o věkovou závislost či spíše o vliv herecké role. Celkově výzkum prozodických vlastností u filmových záznamů je nevhodný. I když v metodách rozpoznávání mohou i herci figurovat a bylo by dobré nalézt vhodná učící data (trend prozodických vlastností).

V tab. 7.8 jsou uvedené dosažené výsledky v této práci.

Parametr	Trend
F0	Klesá
Jitter , shimmer	Roste
VOT (okluze/exploze)	Klesá / klesá
F1 /a/, /e/, /i/, /o/, /u/	Klesá
F2 vokálů /a/, /e/, /o/	Klesá
F2 vokálů /i/, /u/	Roste
VSA	Klesá
VAI	Roste
F2i/F2u	Klesá
Rychlost řeči	Roste v explozivě /d'/

Tab. 7.8: Výsledky věkové závislosti akustických parametrů

8. Závěr

Práce se skládá z několika částí: rešerše, vzniku lidského hlasu, softwaru pro zpracování řečových signálů, popisu vybraných akustických parametrů, které byly dále experimentálně ověřeny na vytvořené databázi dat.

Kapitola 3 byla věnována analýze zkoumané problematiky, analýze dosavadních studií. Zpracovány byly podélné i příčné studie, shrnutí jejich použité metodiky, databázi dat, získaných trendů s rostoucím věkem, tab. 3.1.1 a 3.2.1.

Na základě zpracovaných výsledků studií byla navržena metodika ověřování věkové závislosti vybraných akustických parametrů. Část byla věnována práci zabývající se českými explozívy [15]. Tato práce byla přínosná pro segmentaci explozívy na jednotlivé fáze.

Kapitola 4 byla věnována vzniku lidského hlasu od prvotního signálu vycházejícího z hlasivek po diferencující se v rezonančních dutinách. Dále byla věnována podkapitola samotné artikulaci, jak jsou rozděleny samohlásky i souhlásky. Bylo zpracované rozdělení samohlásek podle prvních dvou formantů a jejich pozice v Hellwagovu schématu. Souhlásky byly také rozděleny a detailně popsány byly explozívy. Důležité rozdíly mezi znělými a neznělými plozívy, na základě jejichž vlastností probíhala segmentace v experimentální části.

V kapitole 5 byl popsán software Praat pro zpracování řeči, který byl používán pro analýzu hlasového signálu jak v několika studiích, tak v této práci.

V kapitole 6 byly popsány jednotlivé vybrané akustické parametry, rozdělené podle třech základních vlastností řeči (fonačních, artikulačních a prozodických). Detailněji byly popsány základní frekvence F_0 , frekvenční variabilita jitter, amplitudová variabilita shimmer, formantové frekvence, velikost vokální plochy VSA, příznaky pro rozpoznávání Parkinsonovy choroby (artikulační index VAI, poměr F_{2i}/F_{2u}), počátek znělosti VOT s popisem segmentace explozívy na jednotlivé fáze, rychlost řeči a podíl pauz v promluvě.

V kapitole 7 bylo zpracováno experimentální ověřování výše zmíněných akustických parametrů. Byla vytvořena databáze filmových záznamů. Formát dat byl upraven pro další zpracování z avi na wav. U každého analyzovaného parametru je uvedena metodika zkoumání, jaká část záznamů byla pro výzkum použita a jsou uvedené výsledky experimentu. Největší problém byl při zkoumání prozodických vlastností řeči, rychlosti řeči a podílu pauz v promluvě. Filmové záznamy nebyly vhodným materiálem k ověřování těchto charakteristik, protože záznamy pro dané roky byly omezené hereckou rolí.

V kapitole 9 je diskutováno nad možnými cíly a postupy dalšího zkoumání věkové závislosti změn hlasových charakteristik, které jsou důležité z hlediska včasného rozpoznání změn, které nesouvisí s fyziologickým stárnutím zdravého hlasu. Dále jsou důležité charakteristiky pro práci metod rozpoznávání osob, aplikace ovládané hlasem a jiné komerční aplikace využívající poznatky o změnách lidského hlasu v průběhu celého života.

9. Naplnění cílů, perspektivy dalšího řešení

V rámci práce byla navržena metodika ověřování věkové závislosti vybraných akustických charakteristik na databázi dat. V současné době je stále nedostatek podélných studií zabývajících se hlasem (mužským, ženským) v průběhu celého života. Filmové záznamy herců jsou vhodným materiálem pro vytváření rozsáhlé databáze záznamů hlasu, na kterých se může analyzovat chování hlasu.

Mým cílem bylo vytvořit databázi již nastříhaných záznamů s popisem, který slouží pro rychlejší orientaci v datech. Lze díky němu navrhnout metodiku zkoumání vybraných částí. Nejvíce času zabralo stříhání promluv a následný poslech všech záznamů dokola, hledání prodloužených fonací, či slov pro analýzu počátku znělosti. S popisem může kdokoli navázat v pokračování analýzy a nejen v záznamech od herečky Ivany Chýlkové. Je třeba vytvářet databáze s popisem u všech hlasů, aby byla snadnější práce se zvukovým materiálem a člověk neztrácel čas hledáním určitých slov/slabik/hlásek. V této práci byl trochu problém se zaměřením některých filmů. Většina záznamů byla ovlivněna emocemi a to ztěžovalo ověřování některých charakteristik. Herec také pracuje s hlasem díky roli, dokáže hýbat s frekvencemi základního hlasu ve velkém rozsahu, a tak může docházet k chybným výsledkům. Hodnocení akustického chování hlasu je založeno na subjektivním poslechu dat. Každý může mít jiný pohled na hranici určitého zvuku, proto nejsou nejdůležitější změny číselných hodnot, ale trend daného parametru s rostoucím věkem.

Díky zkoumání možností softwaru použitého pro zpracování řečových signálů byly získány hodně pěkné výsledky v oblasti formantových frekvencí. Je třeba program nastavovat různě a neříkat se jen na výsledky vypočtené programem, ale pozorovat i změny ve spektru.

V práci byla na základě rešerše počátku znělosti nejen studií věkové závislosti, ale i vlastností explozí (českých) popsána metodika pro segmentaci akustického signálu. Řada záznamů byla horší kvality, ale i přesto se podařilo díky znalostem vlastností jednotlivých fází dané segmenty rozpoznat. Samozřejmě je to z velké části subjektivní hodnocení.

Do budoucna by bylo dobré pokračovat ve zkoumání počátku znělosti u dalších explozí. Dále se věnovat zejména prozodickým vlastnostem řeči, rychlosti řeči, která souvisí i s počátkem znělosti, ale i délkám pauz. Chtělo by to rozšířit databázi o další hlasy, ženské pro srovnání s výsledky z této práce, ale i o mužské na základě stejné metodiky. Snažit se vybrat dostatečně kvalitní hlasové záznamy. V této práci byly použité nahrávky stažené z volně dostupných filmových databází i v horší kvalitě nahrávání a přesto bylo dosaženo uspokojivých výsledků. Dále se nepovedl potvrdit větší frekvenční rozsah u mladších jedinců. Pro jeho analýzu je třeba získat záznamy bez větších vlivů emocí a hereckého nasazení, většího důrazu.

V České republice se zabývá řada specialistů výzkumem tvorby jednotlivých hlásek (Z. Palková, P. Machač, M. Krčmová). Na základě jejich poznatků o vlastnostech jednotlivých hlásek lze postavit metodiku zkoumání změn českého hlasu.

Na základě výsledků trendů akustických charakteristik by se mohly vytvořit modely, které by se učily na datech zdravých jedinců a mohly tak predikovat chování s rostoucím věkem. Rozvíjeli by se samotné metody rozpoznávání osob a včasná detekce poruch hlasu.

10. Seznam použité literatury

[1] *Decoster, W., Debruyne, F.: Longitudinal Voice Changes: Facts and Interpretation.*

Journal of Voice, pp. 184-193, 2000.

[2] *Brückl, M.: Women's Vocal Aging: A Longitudinal Approach.* Interspeech, 4, pp. 2856-2859, 2007

[3] *Harrington, J., Palethorpe, S., Watson, C.I.: Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers.* Interspeech, 2, pp. 1081-1084, 2007

[4] *Xue, S.A., Hao, G.J., Xu, L., Moranski, T.: Frequency Changes in Women Over Time.*

Asia Pacific Journal of Speech, Language, and Hearing, 11 (3), pp. 189-194, 2008.

[5] *Reubald, U., Harrington, J., Kleber, F.: Vocal aging effects on F0 and the first formant: A longitudinal analysis in adult speakers.* Speech Communication, 52 (7-8), pp. 638-651, 2010.

[6] *Xue, S.A., Deliyiski, D.: Effects of aging acoustic voice parameters: Preliminary normative data and educational implications.* Educational Gerontology, pp. 159-168, 2001.

[7] *Schötz, S., Müller, C.: A Study of Acoustic Correlates of Speaker Age.* Lecture Notes in Computer Science, 4441 LNAI, pp. 1-9, 2007

[8] *Schötz, S.: Acoustic Analysis of Adult Speaker Age.* In Speaker Classification, C. Müller, Ed., vol. 4343 of Lecture Notes in Computer Science / Artificial Intelligence. Springer, Heidelberg - Berlin - New York, 2007, this issue.

[9] *Torre III, P., Barlow, J.A.: Age-related changes in acoustic characteristics of adult speech.* Journal of Communication Disorders, 42 (5), pp. 324-333, 2009.

[10] *Vipperla, R., Renals, S., Frankel, J.: Ageing voices: The effect of changes in voice parameters on ASR performance.* Eurasip Journal on Audio, Speech, and Music Processing, art. no. 525783, 2010.

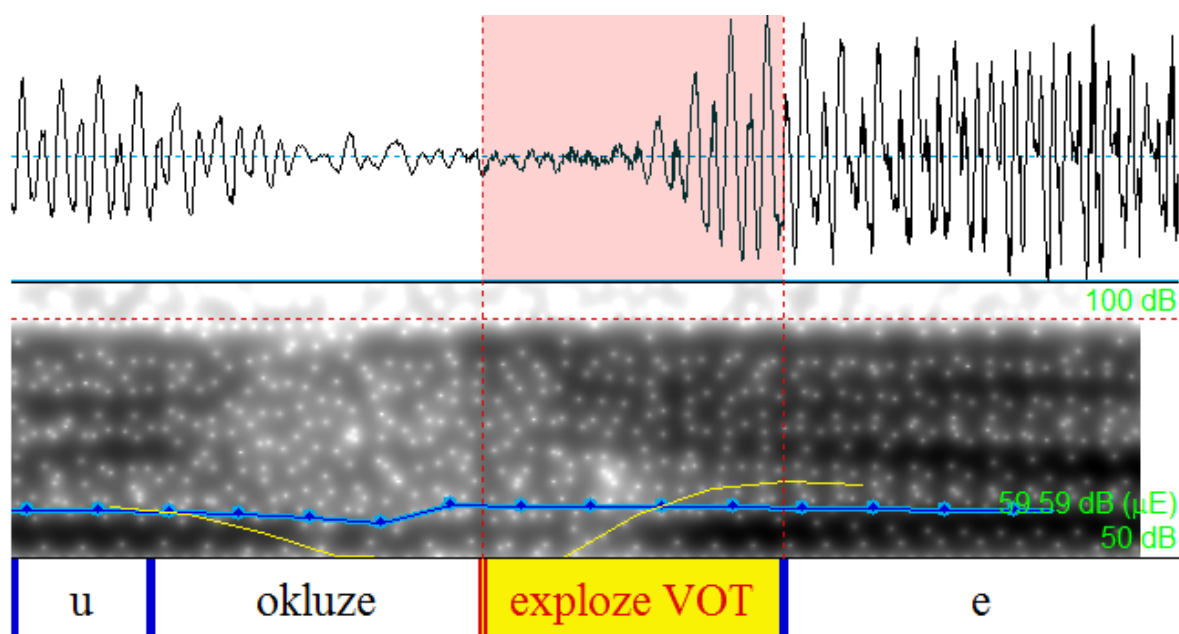
[11] *Markó, A., Bóna, J.: Fundamental frequency patterns: The factors of age and speech type.* Proceedings of the Workshop „Sociophonetics, at the crossroads of speech variation, processing and communication“, 2010.

[12] *Das, B., Mandal, S., Mitra, P., Basu, A.: Effect of aging on speech features and phoneme recognition: a study on Bengali voicing vowels.* Springer Science+Business Media, 2012.

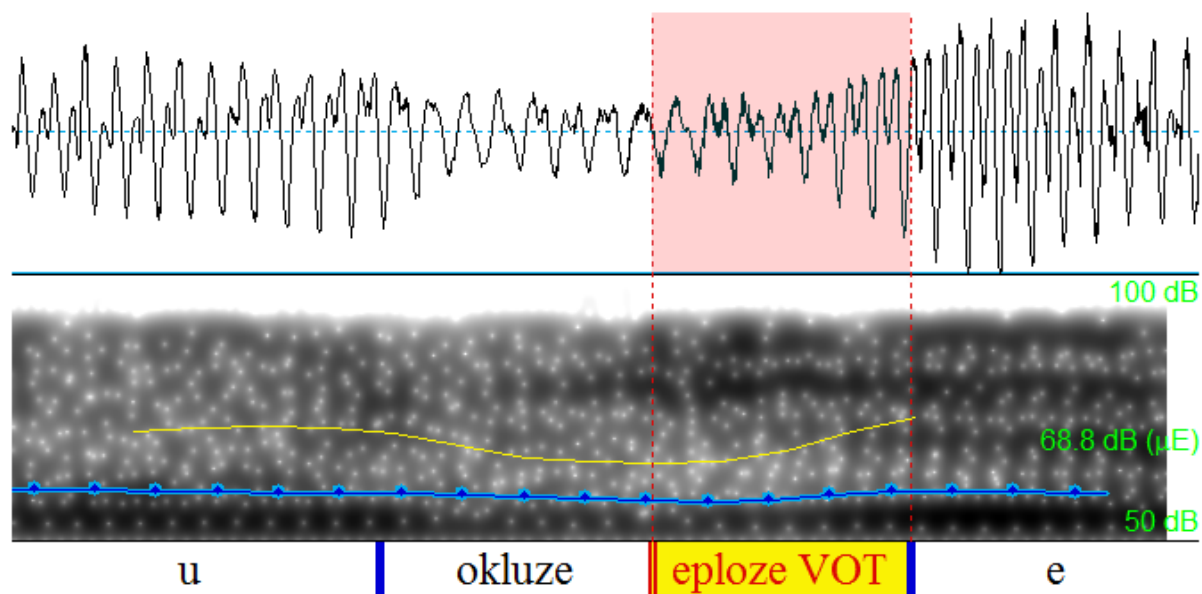
- [13] *Báňa, J.*: **Porovnání analýzy řečového signálu v závislosti na věku a pohlaví mluvčího.** Bakalářská práce, VUT v Brně, 2008
- [14] *Košturiaková, L.*: **Věková závislost lidské řeči.** Diplomová práce, ČVUT v Praze, 2013.
- [15] *Machač, P.*: **Temporální a spektrální struktura českých explozív.** Disertační práce, Univerzita Karlova v Praze, 2006.
- [16] *Janda, J.*: **Posuzování logopedického věku dítěte.** Teze k disertační práci, ČVUT v Praze, 2012.
- [17] *Uhlíř, J., Sovka, P., Pollák, P., Hanžl, V., Čmejla, R.*: **Technologie hlasových komunikací.** Nakladatelství ČVUT, Praha, 2007.
- [18] *Psutka, J., Müller, L., Matoušek, J., Radová, V.*: **Mluvíme s počítačem česky.** Nakladatelství Academia, Praha, 2006.
- [19] *Krčmová, M.*: **Fonetika a fonologie.** Třetí vydání, Masarykova univerzita, 2009.
URL: http://is.muni.cz/do/1499/el/estud/ff/ps09/fonetika/tisk_2009/index.html
- [20] *Machač, P., Skarnitzl, R.*: **Fonetická segmentace hlásek.** Nakladatelství Epoque, Praha, 2009.
- [21] URL: http://kaleidoskop.upol.cz/old/kal2010/jak_vznika_lidsky_hlas.pdf
- [22] URL: <http://savethevowels.org/praat/UsingPraatforLinguisticResearchLatest.pdf>
- [23] *Karásek, M.*: **Diagnóza Parkinsonovy choroby z řečového signálu.** Diplomová práce, VUT v Brně, 2011.
- [24] *Bílý, O.*: **Moderní řečové příznaky používané při diagnóze chorob.** Diplomová práce, VUT v Brně, 2010.

Příloha A – obrazová příloha

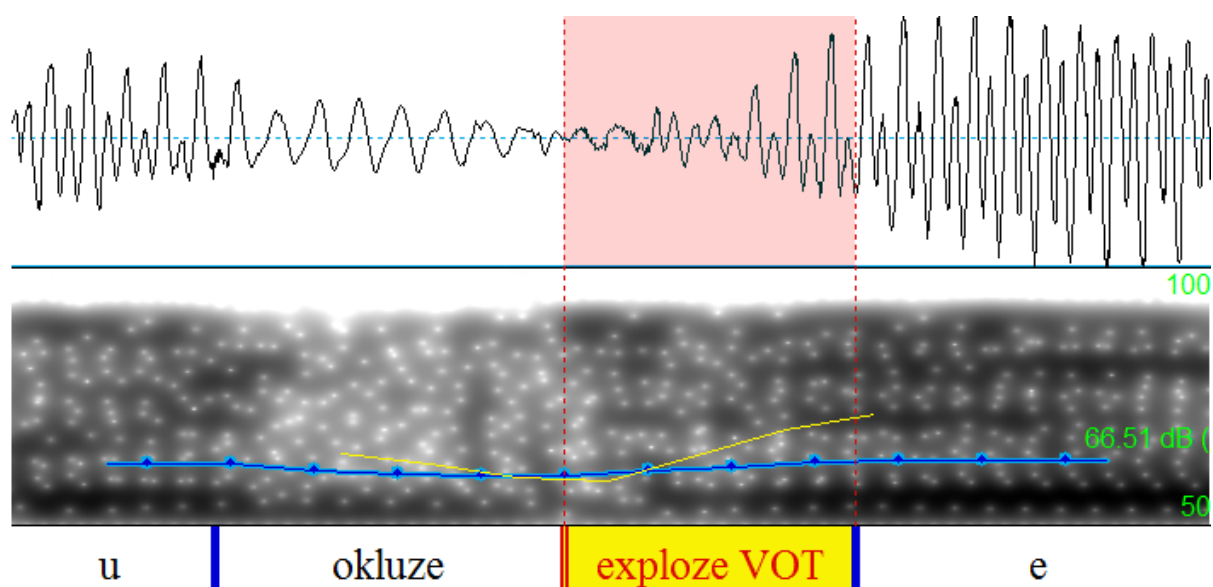
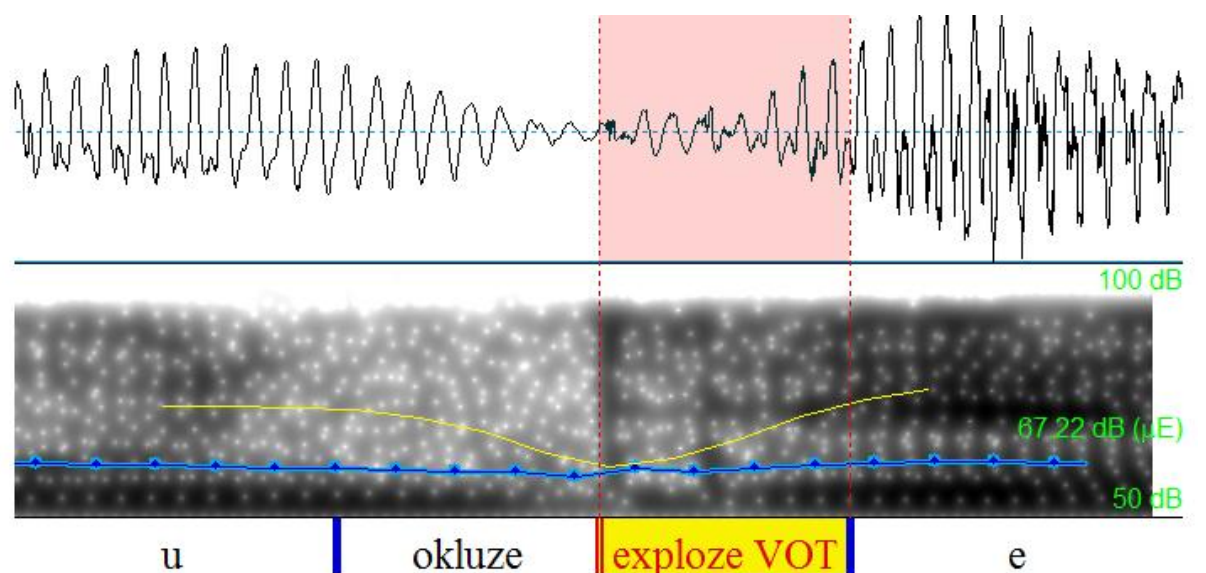
POČÁTEK ZNĚLOSTI pro jednotlivé roky



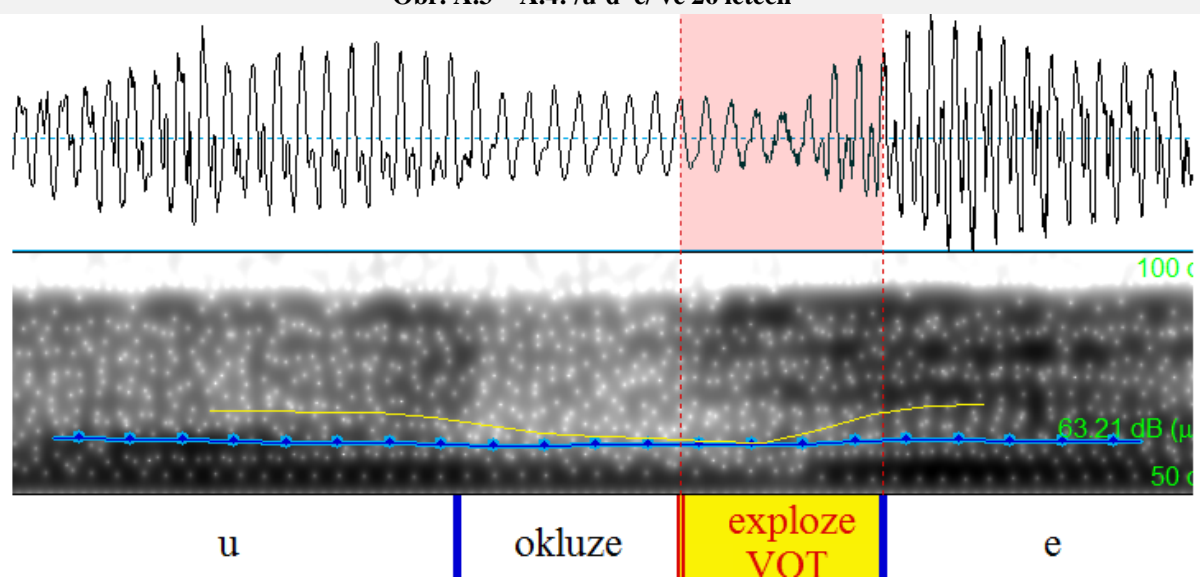
Obr. A.1: /u-d'-e/ ve 22 letech



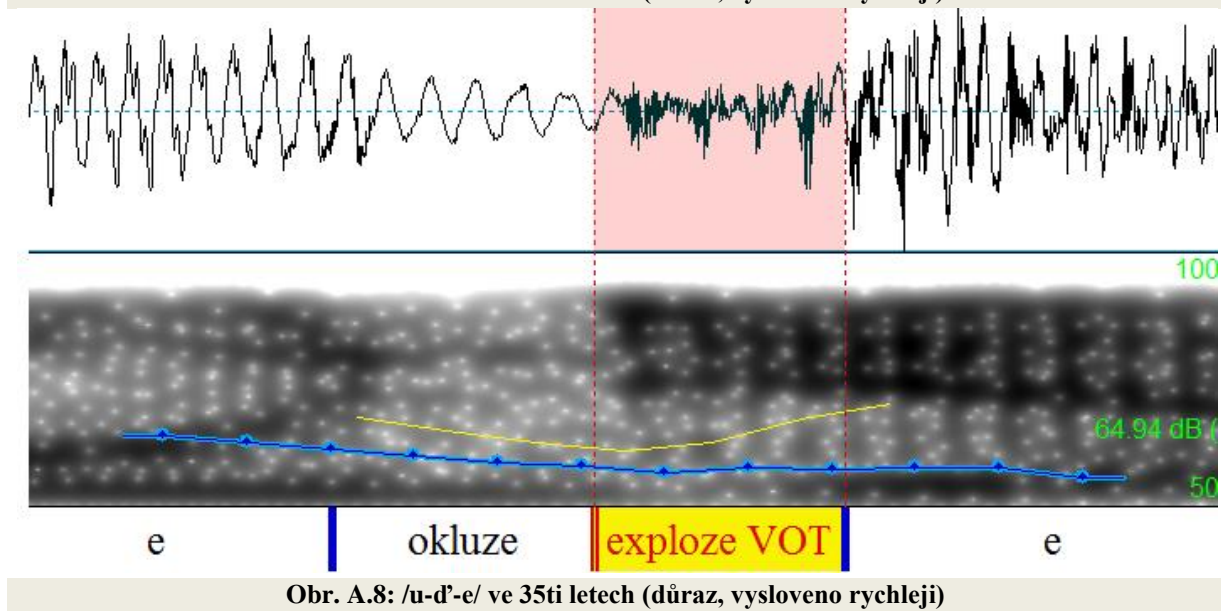
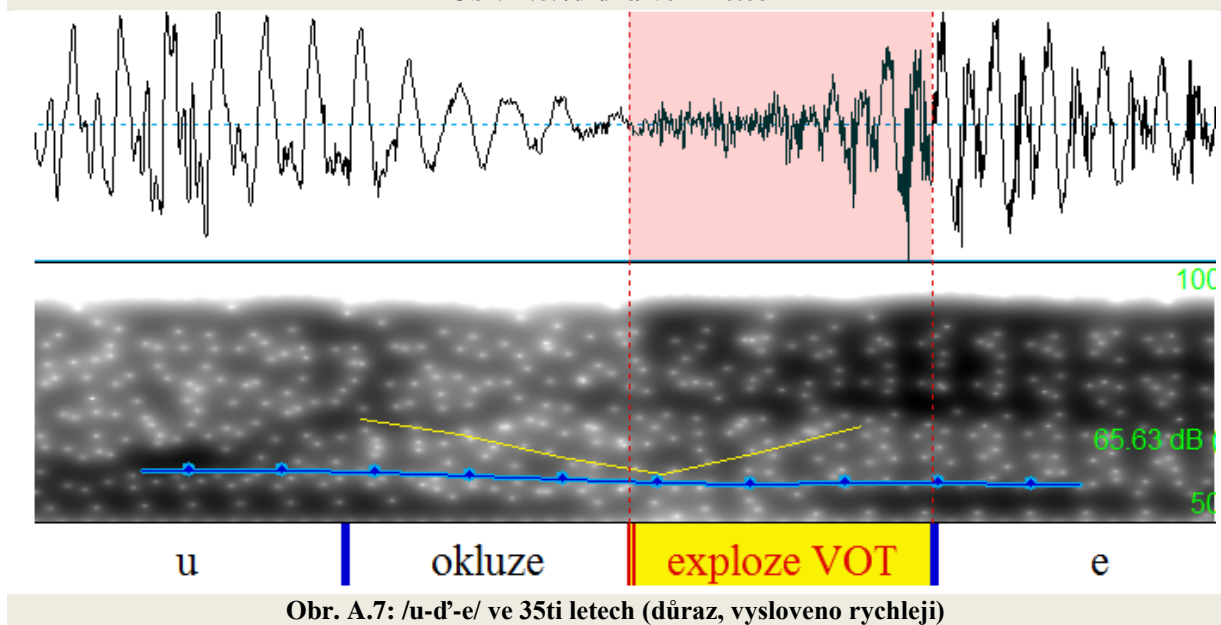
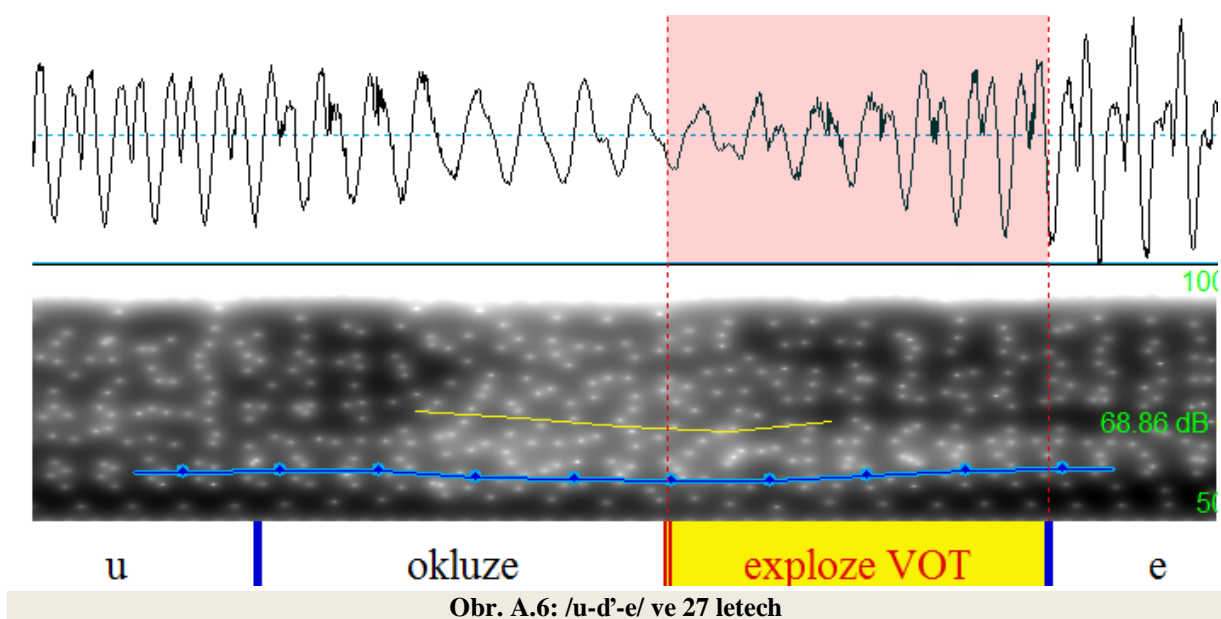
Obr. A.2: /u-d'-e/ ve 23 letech

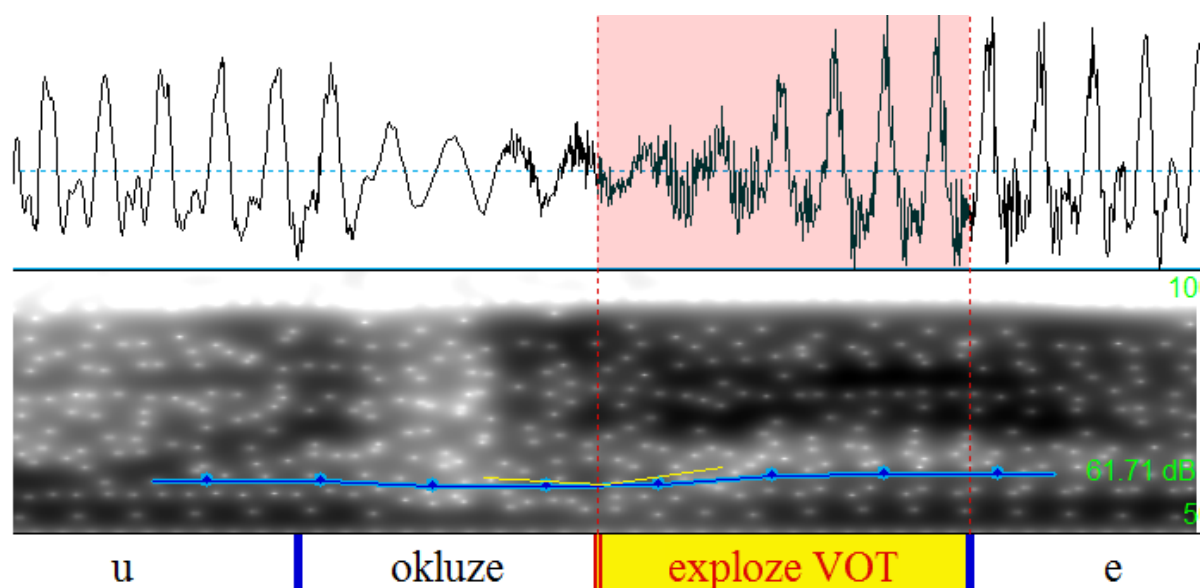


Obr. A.3 – A.4: /u-dʰ-e/ ve 26 letech

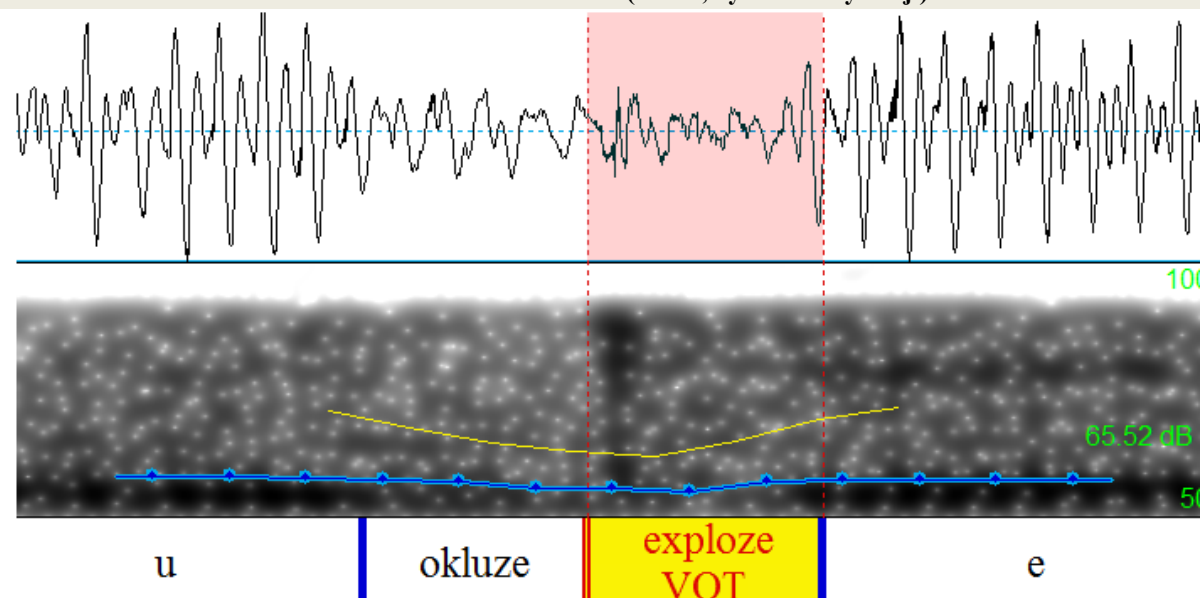


Obr. A.5: /u-dʰ-e/ ve 26 letech

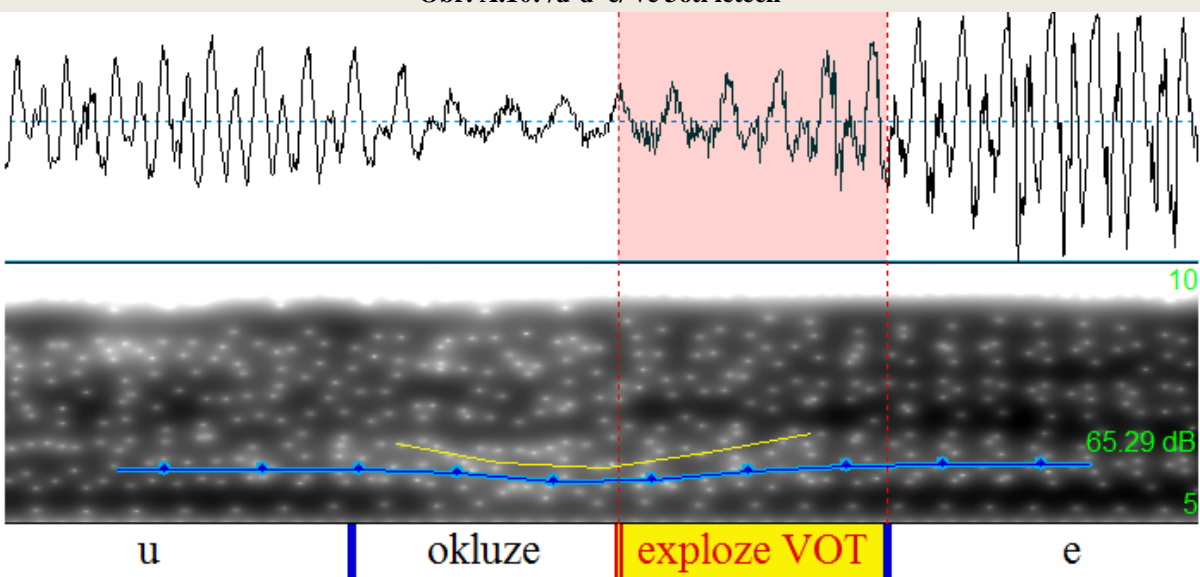




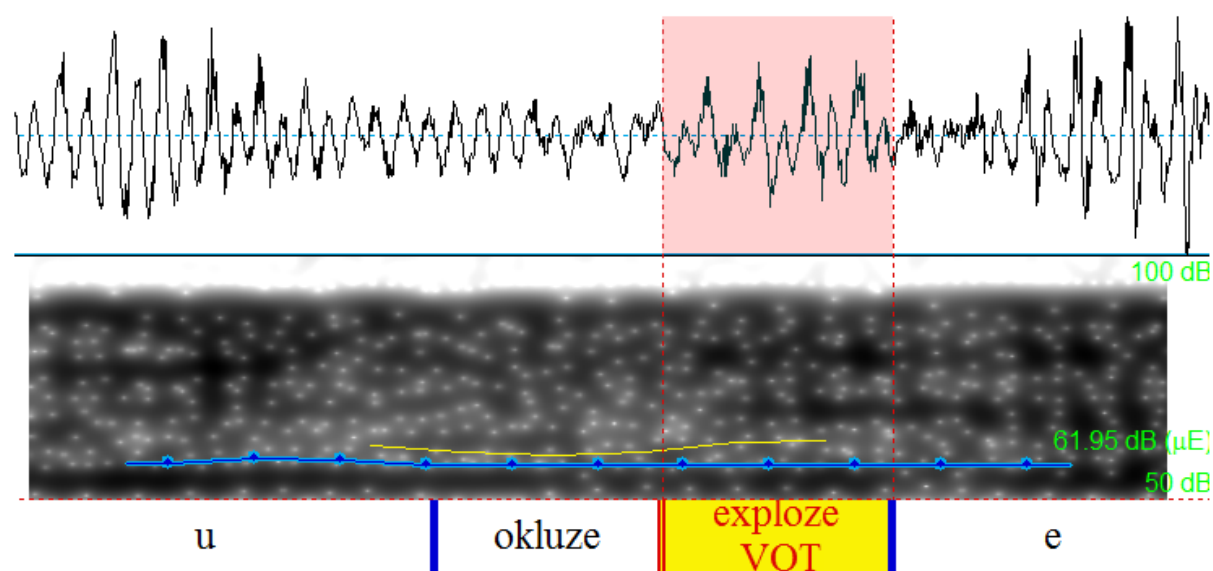
Obr. A.9: /u-d'-e/ ve 35ti letech (důraz, vysloveno rychleji)



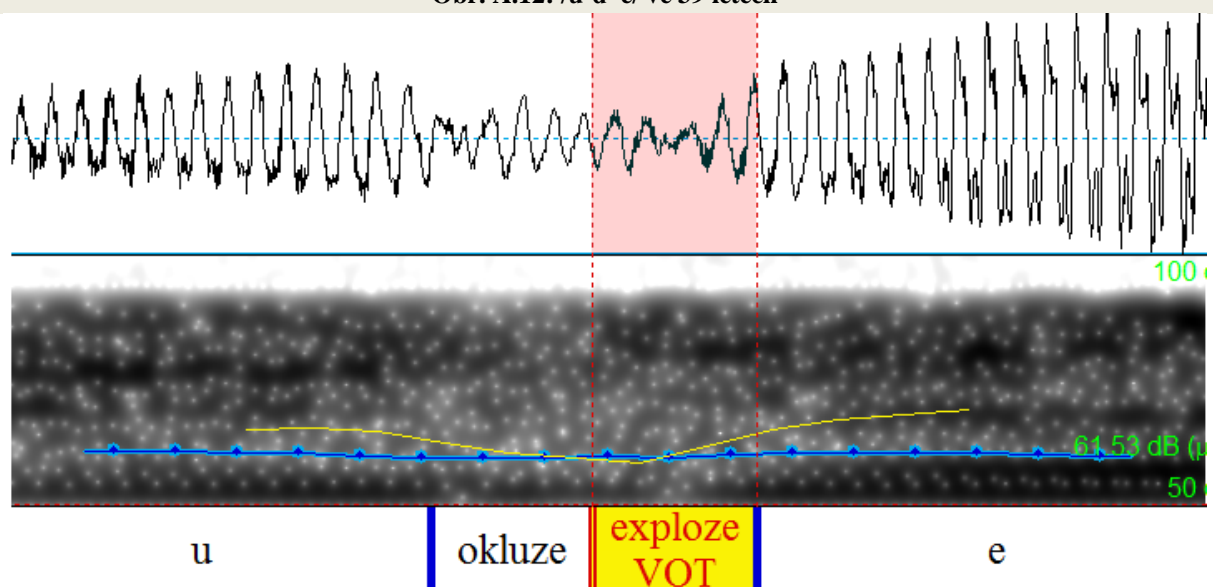
Obr. A.10: /u-d'-e/ ve 36ti letech



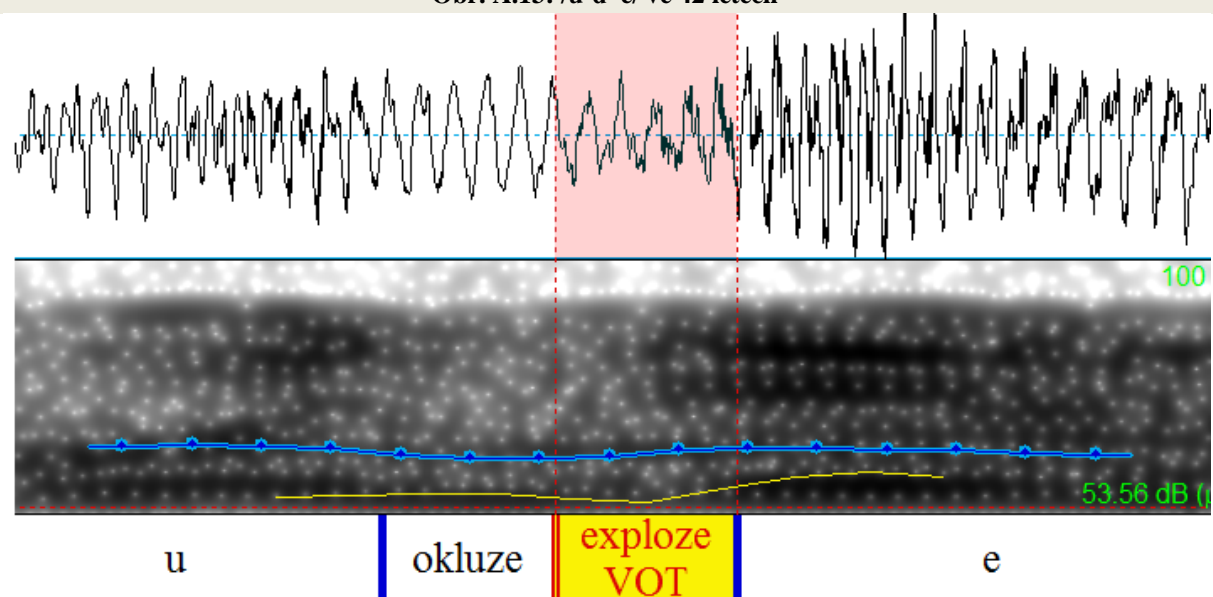
Obr. A.11: /u-d'-e/ ve 37 letech



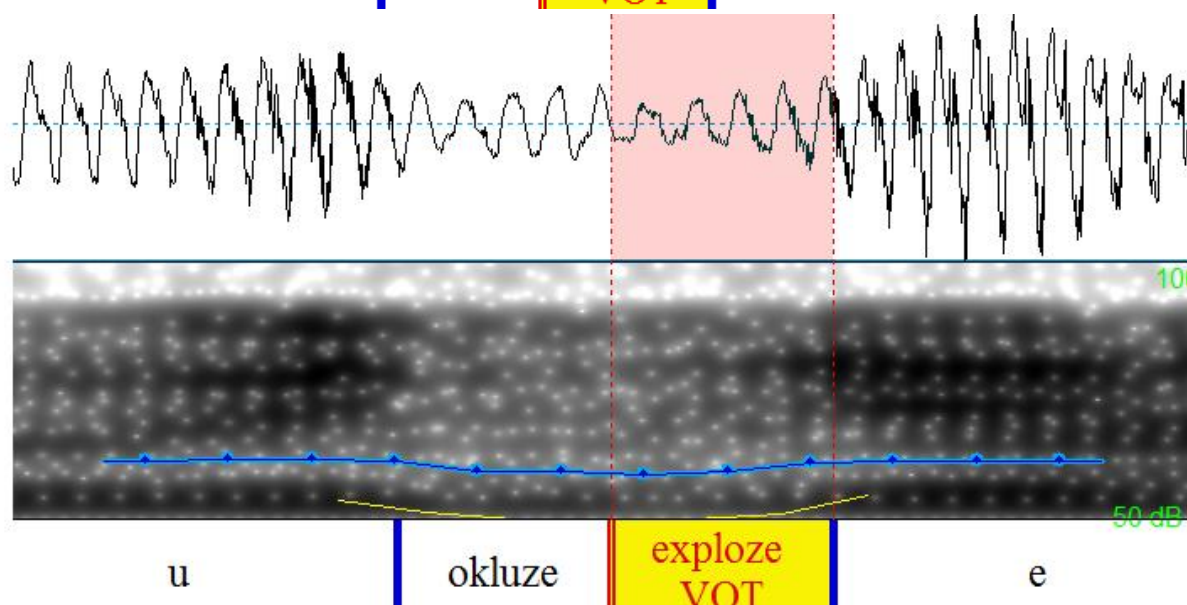
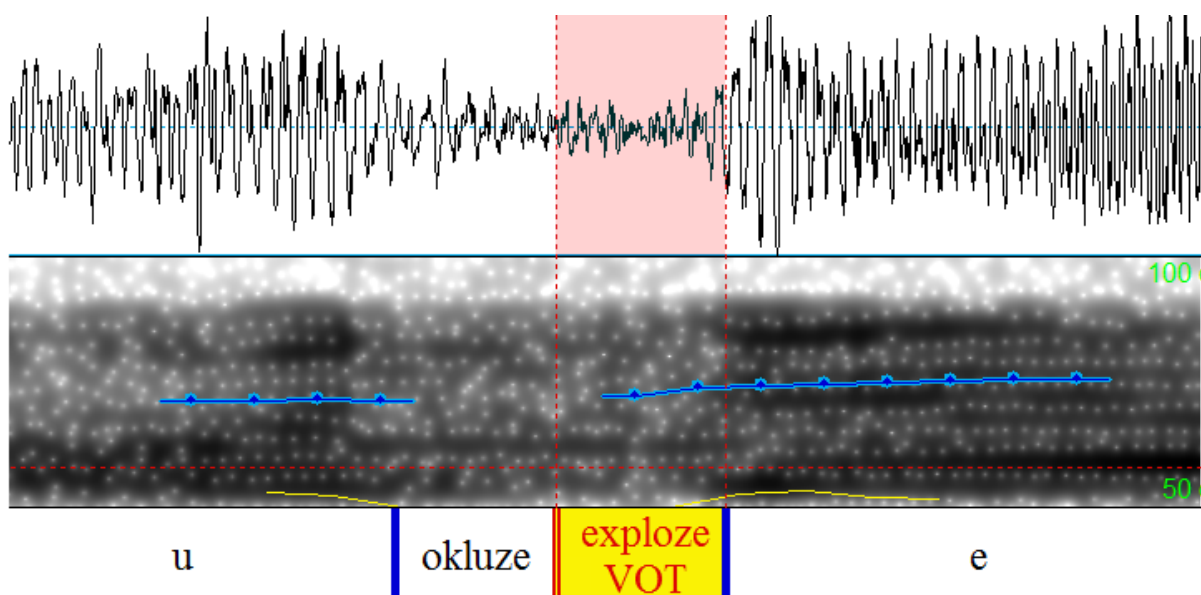
Obr. A.12: /u-dʰ-e/ ve 39 letech



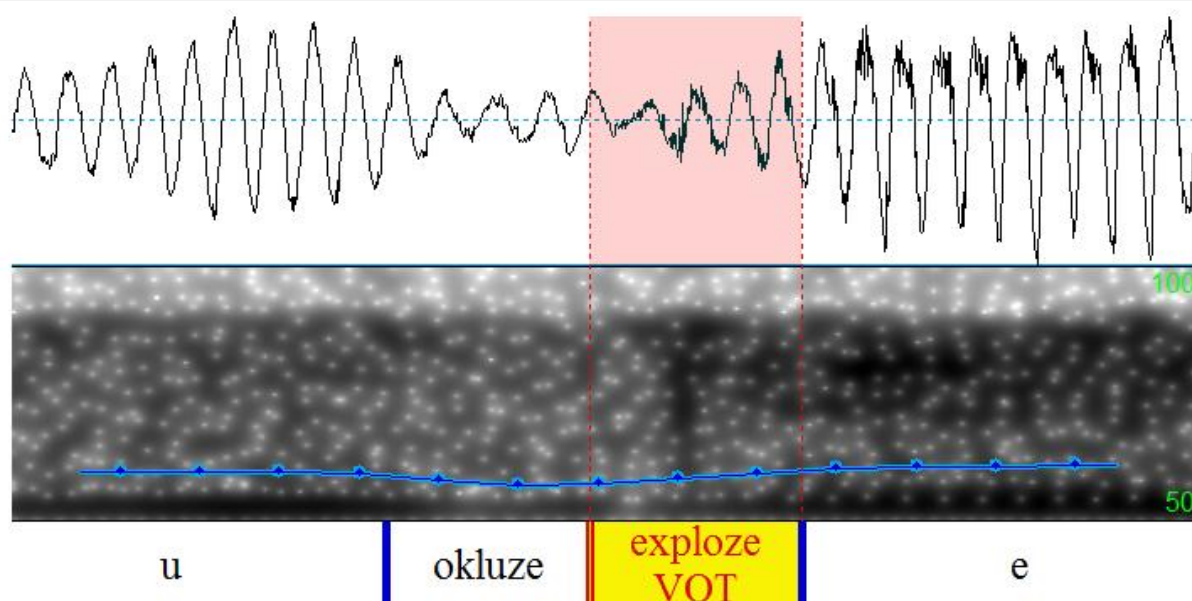
Obr. A.13: /u-dʰ-e/ ve 42 letech



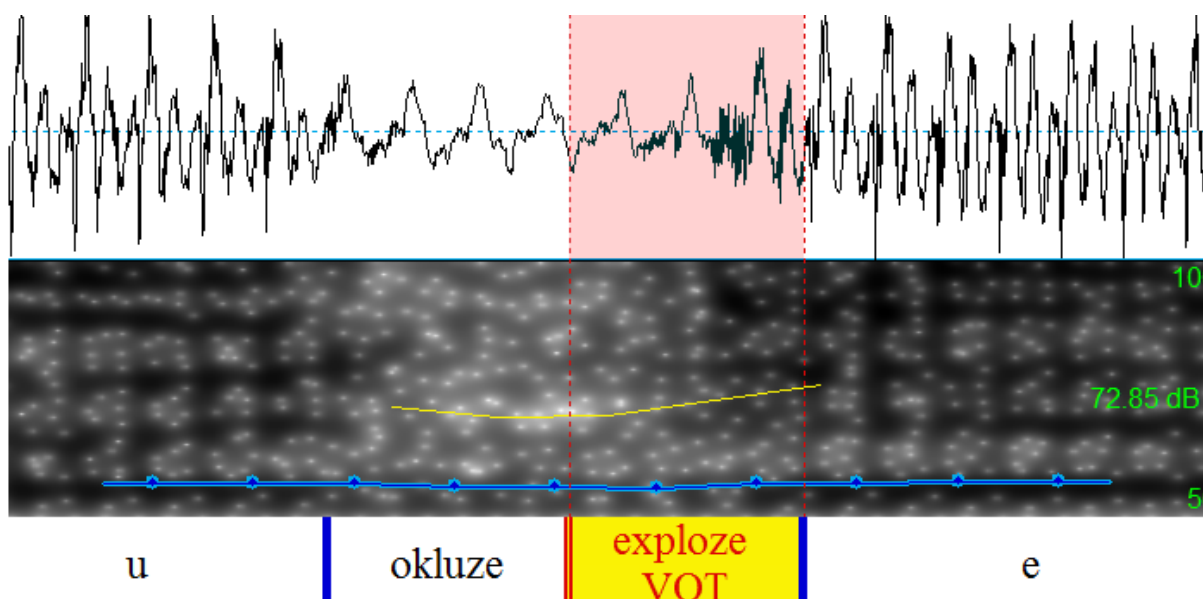
Obr. A.14: /u-dʰ-e/ ve 48 letech



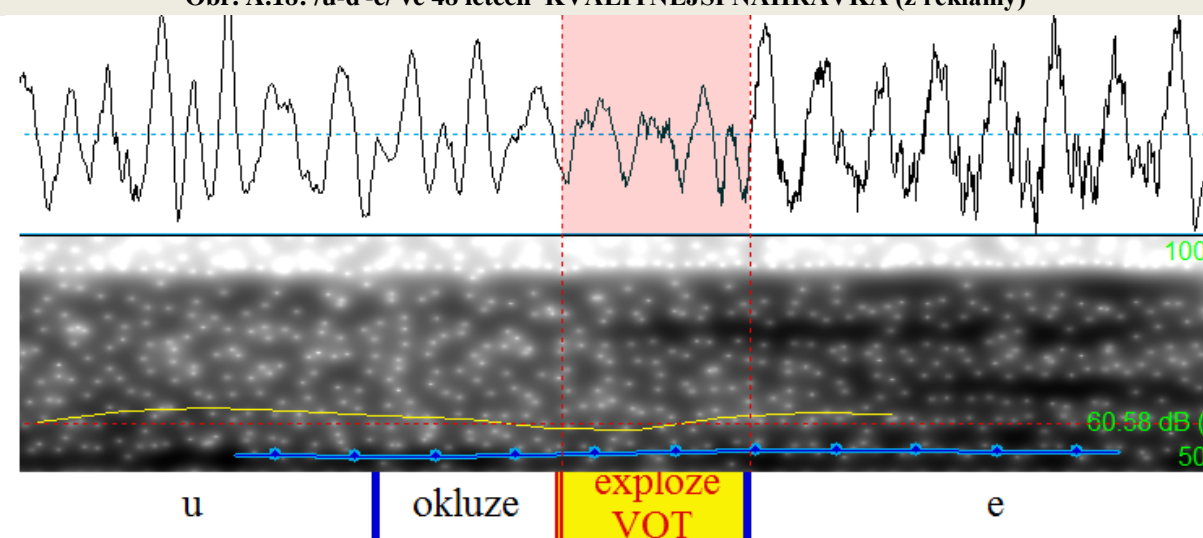
Obr. A.15-A.16: /u-dʰ-e/ ve 48 letech



Obr. A.17: /u-dʰ-e/ ve 48 letech



Obr. A.18: /u-dʰ-e/ ve 48 letech KVALITNĚJŠÍ NAHRÁVKA (z reklamy)



Obr. A.19: /u-dʰ-e/ v 50ti letech horší kvalita (špatně šlo rozeznat /u/)

Příloha B – tabulková příloha

vek [roky]	mean F0 /a/ [Hz]	vek [roky]	mean F0 /a/ [Hz]
22	262,74	37	201,01
23	249,68	38	200,86
24	240,45	39	200,27
26	237,18	42	199,22
27	236,37	44	196,54
29	227,06	45	180,96
31	225,35	46	177,91
35	225,2	48	170,83
36	205,06	50	169,29

Tab. B1: Průměrné hodnoty F0 naměřených na prodloužené fonaci vokálu /a/

vek [roky]	jitter [%]	shimmer [%]	vek [roky]	jitter [%]	shimmer [%]
22	0,92	5,48	37	1,58	11,26
23	1,07	6,3	38	1,61	12,06
24	1,1	7,28	39	1,61	12,06
25	1,11	7,35	42	1,61	12,14
26	1,15	8,61	44	1,62	12,16
27	1,16	8,61	45	1,7	12,17
29	1,16	8,61	46	1,7	12,63
30	1,16	8,71	48	1,71	12,65
31	1,21	9,05	49	1,88	12,78
34	1,21	10,15	50	2,36	12,83
35	1,23	10,22	51	2,36	14,42
36	1,41	11,21			

Tab. B2: Průměrné hodnoty jitteru a shimmeru pro jednotlivé roky.

	A		E		I		O		U	
	F1 [Hz]	F2 [Hz]	F1 [Hz]	F2 [Hz]	F1 [Hz]	F2 [Hz]	F1 [Hz]	F2 [Hz]	F1 [Hz]	F2 [Hz]
22	1039	1513	711	2066	462	2559	700	1118	412	614
24	1038	1511	702	2065	461	2559	700	1117	412	618
25	1035	1511	701	2065	461	2561	699	1117	412	620
26	1034	1499	698	2060	458	2570	691	1117	412	622
27	1033	1499	695	2060	455	2598	689	1106	410	622
29	1024	1499	677	2060	450	2621	673	1106	410	625
30	1005	1496	675	2044	441	2635	671	1106	409	627
31	933	1489	674	2040	392	2651	646	1105	406	641
34	917	1482	669	2021	384	2672	644	1105	404	647
35	870	1470	660	2017	370	2681	642	1105	394	650
36	863	1469	654	2016	363	2692	640	1103	363	650
37	856	1468	651	2012	362	2704	627	1064	355	655
38	853	1463	645	2011	361	2705	625	1063	354	657
39	849	1450	643	2010	351	2705	624	1063	353	665
42	841	1448	639	2010	343	2705	623	1063	346	667
44	840	1448	633	2009	342	2707	621	1057	340	688
45	835	1448	589	1987	325	2707	573	1014	327	690
46	794	1396	571	1981	325	2709	567	1014	326	690
48	777	1380	571	1969	322	2726	566	1000	325	691
49	756	1361	557	1960	321	2729	545	998	313	694
50	755	1350	556	1959	314	2733	543	996	304	695
51	730	1350	523	1957	304	2769	522	989	303	701

Tab. B3: Průměrné hodnoty F1 a F2 pro všechny vokály /a/, /e/, /i/, /o/, /u/

věk	okluze [ms]	exploze [ms]	celkem [ms]
22	41,32	37,86	79,18
23	39,29	37,22	76,51
26	38,08	34,04	72,12
27	37,03	34,28	71,31
35	26,03	28,03	54,06
36	25,78	27,18	52,96
37	24,21	24,43	48,64
39	23,39	23,77	47,16
42	23,11	23,62	46,73
48	22,17	22,84	45,046
50	20,1	20,78	40,88

Tab. B4: Průměrné hodnoty dob trvání okluze, exploze a celkové doby trvání explozívy.

Příloha C – seznam filmů a seriálů

Rok	Věk [roky]	Název filmu/seriálů/záznamu
1985	22	Dobré světlo
1986	23	Černá punčocha, Krajina s nábitkem, Můj hříšný muž, Papillio
1987	24	Hauři
1988	25	Tichý společník
1989	26	Čas sluhů
1990	27	Divoká svině, Houpačka, Král kolonád, Ta naše písnička česká, V žáru královské lásky
1992	29	Přítelkyně z domu smutku (2.díl)
1993	30	Sedmero krkavců
1994	31	Díky za každé nové ráno, Vášnivý polibek
1997	34	Stůj, nebo se netrefím
1998	35	Čas dluhů, Jezerní královna
1999	36	Ze života pubescentky
2000	37	Oběti a vrazi
2001	38	Černí andělé (5.díl Experiment)
2002	39	Musím tě svést
2005	42	Oběti 5 – Hadí tanec
2007	44	Cesta do Vídně a zpátky
2008	45	Brainstorm
2009	46	Klub osamělých srdcí
2011	48	Perfect Days – I ženy mají své dny, Holit nebo neholit (reklama)
2012	49	Stín smrtihlava
2013	50	Donšajni
2014	51	To na tobě doschne s Ivanou Chýlkovou (ukázka – čtení ve školce)

Tab. C1: Databáze filmů ze kterých bylo získáno 962 vět.

Příloha D – obsah přiloženého CD

K této práci je přiložené CD, na kterém je uložena diplomová práce v elektronické podobě a naměřená data.

- diplomova_prace_behenska.pdf
- seznam_vet.pdf - celkový seznam vět získaných ze filmů z přílohy C
- Praat.Collection – segmenty /u – d' – e/ v Praatu
- Praat.Collection2 – segmentovy /e – d' – e/ v Praatu
- f0_a.xls – naměřené hodnoty základní frekvence
- formanty.zip – naměřené hodnoty formantů